UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE
EN
MATHÉMATIQUES ET INFORMATIQUE APPLIQUÉES

PAR

LAHYANE ADIL

Vérification des signatures manuscrites

Département de mathématiques et d'informatique
Université de Québec à Trois-Rivières

Février 2002

REMERCIEMENTS

Je tiens à remercier très sincèrement tous ceux qui m'ont aidé à l'accomplissement de ce travail.

Mes remerciements sont formulés aux membres du Laboratoire Interdisciplinaire de Recherche en Imagerie et Calcul Scientifique (LIRICS) tout particulièrement à mes deux directeurs de recherche, Fathallah Nouboud et Alain Chalifour, pour leur soutien considérable et leurs contributions dans l'évolution de ce mémoire.

J'aimerais aussi remercier Louis Paquette professeur, d'informatique à l'université du Québec à Trois-Rivières, pour son aide très bénéfique pour la partie touchant à la programmation et à l'implémentation des méthodes mise au point dans ce mémoire.

Je tiens à remercier les autres membres du comité de lecture, pour leur lecture critique et les commentaires apportés à mon travail de recherche.

Merci à lise Branchaud, secrétaire du programme de maîtrise, pour sa disponibilité permanente.

Je remercie aussi Mostafa Elyassa, professeur de mathématiques à l'université Ibn Zohr, Agadir-Maroc, pour son aide considérable dans le domaine de la prétopologie mathématique.

Merci du plus profond de mon cœur, à ma famille, ma mère Latifa, mon père Lahcen, mes sœurs Safaa et Hind, mes tantes Sadiia, Khadija et Laila ainsi qu'à ma fiancée Ilham, pour leur amour et leur encouragement, leur support et pour avoir cru en moi tout au long de mes études.

Je tiens finalement à remercier tous mes collègues et amis qui, de prés ou de loin, de part leur soutien moral ont contribué à l'aboutissement de ce projet.

Dédicace

Ce mémoire est dédié à ma grand-mère Fatima. Sans son amour et son support au fil des années, ce mémoire n'existerait pas.

Lahyane Adil

Février, 2002

Résumé

Ce mémoire traite de la vérification automatique des signatures manuscrites. En effet, la signature est reconnue comme mode de validation associé à l'identité d'une personne. Un système automatique permettrait une vérification rapide, systématique et efficace des signatures, et réduirait de manière significative les risques de contrefaçon.

Dans ce mémoire, nous avons développé deux méthodologies complémentaires. Une première permet de créer un système automatique de reconnaissance du type des signatures manuscrites, alors que la deuxième est une approche d'authentification des signatures manuscrites se basant sur les réseaux de neurones et utilisant des mesures géométriques et une distance prétopologique.

Ce mémoire est constitué de trois étapes. La première est un prétraitement qui consiste à filtrer les images de signatures. Par la suite, les caractéristiques géométriques et les distances prétopologique sont calculées, et finalement un réseau de neurones est utilisé pour la vérification des signatures.

TABLE DES MATIÈRES

LISTE DES FIGURES

Chapitre 1

INTRODUCTION

Il existe actuellement plusieurs dispositifs permettant l'identification d'un individu dans le cadre d'applications ou d'activités où l'accès à des informations ou à des lieux physiques ou virtuels sont restreints soit pour des raisons de sécurité, de confidentialité, de repérage et maintes autres raisons. Il suffit de citer les systèmes informatiques, les systèmes bancaires, les systèmes juridiques, les centres de recherche à accès limité, pour ne nommer que ces derniers. Une méthode fréquemment adoptée est l'utilisation de codes alphanumériques, de cartes d'accès ou de mots de passe, pour l'utilisation d'un réseau informatique ou encore d'un système bancaire. Cependant de tels dispositifs, quoique simples, peuvent être accessibles à d'autres individus (perte, divulgation, falsification, etc.) et ne constituent pas des procédés d'identification propre à un seul individu en soi.

Certaines approches sont basées sur l'utilisation de caractéristiques biométriques des individus telles que la taille, le poids, la voix, la vascularisation de la rétine, les empreintes digitales, les signatures manuscrites, etc. Un avantage de ces caractéristiques est qu'elles sont propres à chaque individu et donc difficiles à dupliquer. Par ailleurs, l'acquisition, le stockage, l'évolution dans le temps de certains paramètres ainsi que leur pouvoir discriminant au sein d'une large population demeurent une contrainte. A titre d'exemple, l'identification d'un individu par l'analyse des séquences d'acides de son ADN est très efficace mais son utilisation nécessite des prélèvements, des tests de laboratoire et demeure onéreuse. Un paramètre fréquemment utilisé, en particulier par les services judiciaires, est la prise d'empreintes digitales, ici encore, le stockage et l'analyse en temps réel d'une empreinte, par exemple pour activer un système d'ouverture de portes avec senseur, demeurent des contraintes.

Par ailleurs, la signature manuscrite d'un individu représente un bon compromis : tout en étant relativement fiable, elle est facile à acquérir et elle est socialement bien acceptée comme mode d'identification. La signature est un moyen utilisé depuis fort longtemps, l'ancêtre étant le sceau, pour authentifier des documents, pour responsabiliser les individus face à des engagements (contrats, etc.). La signature est donc reconnue comme mode de validation associé à l'identité d'une personne.

La signature d'un individu, en tant que tracé, résulte d'un mécanisme complexe propre à celui-ci. On suppose donc, que la signature de chaque individu est unique, qu'elle le caractérise. Évidemment, plusieurs facteurs influent le tracé d'une signature: la position du scripteur, son humeur, sa forme physique. Par conséquent, certaines variations (intra-individu) sont présentes chez une même personne, on cherche donc à caractériser une signature en prenant en compte ces paramètres. La variabilité intra et interpersonnelle, le nombre et les types (aspects culturels) de signatures obligent à la mise au point, et à la recherche, de méthodologies pour vérifier et authentifier les signatures. A cet effet, les systèmes judiciaire et bancaire ont recours, lors de litiges, à des experts pour statuer sur l'authenticité d'un document ou d'une signature. Cette pratique fournit des résultats satisfaisants dans la plupart des cas, mais reste coûteuse en temps et en expertise. En effet, l'authentification d'une signature peut nécessiter plusieurs mois de recherches diverses. Par conséquent, les banques font appel à une telle expertise qu'en présence d'un doute, pour des chèques impliquant des sommes élevées..

Un système automatisé de vérification et d'authentification contournerait, a priori, toutes ces difficultés. La vérification des signatures serait idéalement rapide, systématique et efficace, et réduirait de manière significative les risques de contrefaçon. Plusieurs systèmes ont été développés à ce jour afin d'automatiser la vérification des signatures et on peut diviser ces méthodes en deux classes suivant le mode d'acquisition de l'image des signatures:

- Les systèmes «On-Line» : la signature est obtenue à partir d'une tablette numérique reliée à un ordinateur. Cette méthode permet d'utiliser des informations dynamiques tels que la vitesse, la pression et/ou l'inclinaison du stylo. Ces systèmes sont surtout utilisés pour contrôler l'accès à des zones protégées ou pour vérifier l'identité lors d'une transaction en-ligne (pourvu qu'on ait les dispositifs de saisie adéquat). Ces systèmes ne peuvent pas être utilisés pour vérifier des signatures déjà apposées sur des documents (chèques bancaires par exemple).

- Les systèmes « Off-Line» : la signature est numérisée à partir d'un support physique tel un chèque ou tout autre document. Cependant, ces systèmes permettent de vérifier les signatures à un temps différé. Mais la perte des informations dynamiques rend le processus de vérification plus difficile.

Dans le cadre de ce mémoire, nous aborderons le problème de la vérification des signatures manuscrites à partir d'une saisie « Off-Line», à l'aide d'une approche inspirée des méthodes utilisées par les experts.

Un premier chapitre portera sur la description des différentes approches proposées ces dix dernières années dans le domaine de la vérification des signatures manuscrites, ainsi que sur la description des méthodes utilisées par les experts lors d'une authentification visuelle. Une analyse comparative des performances (rapidité, robustesse, efficacité) est présentée.

Un second chapitre portera essentiellement sur les méthodes d'acquisition et de prétraitement (élimination du bruit) des images numérisées. Les caractéristiques des signatures y seront relevées afin d'identifier les différents jeux de paramètres utilisés dans les méthodes courantes et les paramètres qui retiendront notre attention en regard d'une automatisation d'une approche calquée sur le comportement des experts.

Finalement, nous présentons une nouvelle approche de vérification et d'authentification des signatures basée sur une approche prétopoloique et un dernier chapitre présente les résultats obtenus suite à des tests appliqués sur une base de données (jeu de signatures) pour valider les performances du système de vérification et d'authentification proposé, en regard des performances des méthodes existantes. La performance des algorithmes proposés ainsi que diverses avenues de recherches seront discutées dans la conclusion.

Chapitre 2

Vérification des signatures : définitions et méthodes

Dans ce chapitre nous présentons une bibliographie exhaustive, conduite dans les domaines du judiciaire et des systèmes de vérification automatique de l'identité à partir d'images numériques des signatures manuscrites. Une première section portera sur la démarche des experts lors d'une authentification, puis dans une deuxième section, nous énumérons les différentes approches existantes dans le domaine de la vérification de signatures manuscrites tout en précisant les points forts et les faiblesses de chaque méthode.

2.1 Les signatures manuscrites

Plusieurs facteurs intrinsèques de la signature et de son signataire influencent le processus d'authentification. Pour comprendre la démarche d'un expert, il faut d'abord connaître quels sont ces facteurs et jusqu'à quel point ils peuvent affecter le jugement de l'expert.

2.1.1 Généralités

L'étude de la signature manuscrite est un cas particulier de l'expertise des documents manuscrits. En effet, l'information disponible pour analyser une signature est relativement réduite, comparée aux textes manuscrits. En outre, une signature plus que toute autre forme d'écriture, est le résultat d'un geste spontané et quasi-automatique.

L'habilité du scripteur constitue un élément important de l'authentification. Le scripteur débutant appuie lentement sur le papier, demeure stationnaire avant d'aborder le tracé de la signature et produit des traits irréguliers. Dans le cas du scripteur habile, la plume est en mouvement avant de toucher le papier et produit des traits réguliers. Ainsi l'habilité du scripteur influence la qualité du tracé.

L'instrument d'écriture a lui aussi une influence sur le tracé de la signature. Gayet([1]) a montré que l'utilisation de la plume permet de détecter la présence ou l'absence de mouvement. En effet, une terminaison effilée est signe de mouvement alors qu'un tracé plus épais par endroit est signe d'un temps d'attente. L'expert peut donc suivre sur le tracé l'épuisement progressif de la plume tout au fil de la signature et ainsi apprécier la vitesse de rédaction.

L'arrivée du stylo à bille à rendu cette information impossible à récupérer car il n'y a pas d'écoulement direct de l'encre du réservoir vers le papier, celui-ci adhère à la surface de la bille et y demeure jusqu'à son report sur le papier. L'épaisseur de la ligne n'est pas influencée par un arrêt brusque du stylo, mais elle est plutôt reliée à la variation de la pression appliquée par le scripteur sur le papier. Enfin, un amincissement graduel d'une ligne tracée par une plume-feutre indique que l'outil d'écriture quitte la surface du papier avec une vitesse élevée.

Les variations naturelles dans le tracé sont des caractéristiques intrinsèques de la signature authentique. En effet, deux signatures parfaitement identiques, indiquent que nous sommes probablement en présence d'un faux. Ces variations sont affectées par plusieurs facteurs tels que l'état de santé et émotionnel de l'individu, l'âge et la fréquence de l'écriture ([2],[3],[4]). De plus, une étude effectuée par Evette et Totty ([5]) démontre la variabilité dans les proportions d'une signature. Il est donc nécessaire de toujours isoler les spécimens dans le temps et de comparer les signatures litigieuses avec des authentiques provenant d'une même période.

Plusieurs études dans le domaine des sciences judiciaires ([1],[3]) montrent que le nombre de signatures de références[1] est important. Un nombre adéquat de références se doit de représenter un profil complet des habitudes du scripteur et de ses habilités.

Il est essentiel pour l'expert de considérer l'ensemble des facteurs circonstanciels cités plus haut, lors de la constitution des signatures de référence. De manière générale deux éléments importants devraient toujours être retenus, lors de l'examen d'une signature litigieuse. En premier lieu, il y aura obligatoirement plusieurs points de similarité entre l'incriminée et l'authentique, surtout dans les aspects les plus évidents de la signature comme le tracé des lettres. Deuxièmement, il n'y a pas de réplique exacte entre les authentiques. Donc, la présence de deux signatures qui seraient parfaitement identiques démontre que l'une d'elle est falsifiée.

[1] Nous appelons signature de référence, une signature authentique utilisée comme référence lors de l'étape

2.1.2 Type de signatures

Étant donné les différences d'origine et de culture entre les signataires à travers le monde, on distingue deux types de signatures :

Le type américain

Ces signatures s'apparentent à l'écriture cursive. Il est donc possible, pour un expert qui traite ce type de signatures d'utiliser un texte manuscrit écrit par le signataire pour comparer la forme de l'écriture (figure 2.1).

Le type européen

Ces signatures possèdent une composante graphique importante qui oblige à un traitement global des signatures (figure 2.2).

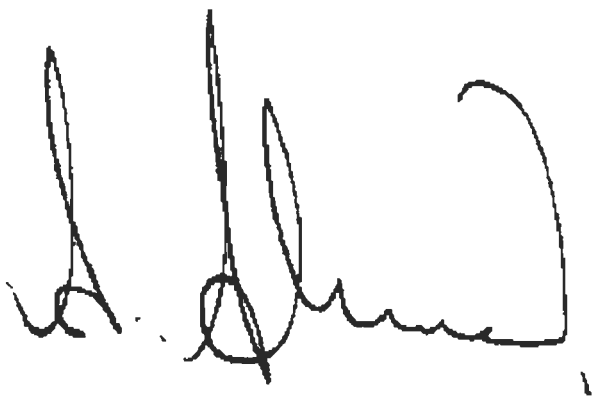

Figure 2.1 : Signature de type américain.

Figure 2.2 : Signature de type européen.

Il existe d'autres types de signatures tel que le type arabe, chinois, etc.., cependant, nous ne traiterons dans ce mémoire que les signatures dites «latines».

2.1.3 Types de faux

Nous consacrons ce paragraphe à l'énumération des types de faux. Suivant le type de faux recherché, l'expert utilise des traitements spécifiques, tel que les proportions de la signature, le nombre de parties des signatures et les projections, dont nous pourrions nous inspirer ultérieurement pour la mise au point d'un système d'authentification automatique (chapitre 4).

Les faux par déguisement

Le faux par déguisement est particulier car il correspond à une signature faite par le signataire présumé (d'origine) mais déguisée délibérément dans le but de pouvoir renier celle-ci ultérieurement. Ces signatures sont généralement ressemblantes malgré une perte d'harmonie dans le tracé et des changements de vitesse. Une étude fine faite par un expert permet de les détecter.

Les faux par imitation servile

Dans le cas d'un faux par imitation servile, le faussaire doit posséder un exemplaire de la signature authentique. Ce faux quoique ressemblant à l'original présente des différences dans les espacements, dans les inclinaisons. De plus le tracé est lent et hésitant d'où des variations visibles de la pression.

Les faux par imitation libre

Pour ce faux le faussaire étudie soigneusement la signature authentique et s'entraîne à la reproduire de mémoire jusqu'à être satisfait du résultat. Ce sont, de l'avis des experts, les faux les plus difficiles à détecter car ils sont très ressemblants et le tracé est spontané. Ils diffèrent des originaux par les proportions relatives des éléments de la signature et par l'alternance des pleins et de déliés.

Les faux par calque

Le faux par calque reproduit fidèlement l'image d'une signature authentique sur un document par l'utilisation d'un moyen de recopiage par exemple : une copie par transparence, avec du carbone, par photocopie. Ce faux est évidemment très difficile à détecter même pour les experts.

Les faux grossiers

Dans le cas d'un faux grossier, le faussaire n'essaie pas de faire un faux ressemblant à un original. Ce faux est fréquent et il est le plus facile à détecter. Ce sont ces faux que nous espérons pouvoir détecter facilement à l'aide de la méthodologie développée dans ce mémoire.

Les faux aléatoires

Les signatures des personnes autres que le signataire présumé sont appelées des faux aléatoires. Pour tester les systèmes, c'est le faux le plus simple à simuler. En effet, il est difficile d'obtenir des banques de faux réels en quantité suffisante alors qu'il suffit de prendre les signatures des autres personnes dans une base de signatures. Les faux aléatoires font partie des faux grossiers et on suppose que les résultats obtenus sur ce type de faux sont identiques à ceux que l'on obtiendrait sur des faux grossiers.

Les faux simples

Le faussaire fabrique une signature à partir du nom du signataire sans imiter un original. Ce faux est généralement peu ressemblant surtout pour les signatures de type européen de nature graphique; il est pourtant plus difficile à détecter que des faux grossiers. Ce type de faux est intéressant d'un point de vue technique (validation) pour les systèmes traitant les signatures de type américain de nature cursive car il permet de tester les systèmes d'authentification avec des faux plus ressemblants que les faux grossiers.

Pour les faux simples, grossiers et aléatoires le tracé est spontané, les caractéristiques pseudo-dynamiques sont donc peu discriminantes, par contre la ressemblance est faible et les caractéristiques sont assez discriminantes. Ces faux sont considérés comme les plus faciles à détecter par un système automatique car ils ne nécessitent pas d'expertise et de plus, on les rencontrent couramment.

Pour les autres faux, leur aspect étant très ressemblant, il est nécessaire d'utiliser des caractéristiques pseudo-dynamiques (trais hésitant, pression, vitesse, pleins et déliés, retouches, arrêts de stylo,...). Ces caractéristiques sont plus ou moins apparentes suivant le stylo utilisé, elles s'avèrent difficiles à extraire automatiquement.

2.2 La démarche expert.

Dans cette section nous présentons deux approches d'analyse graphométrique fréquemment utilisées par les experts. Ces approches exploitent des caractéristiques spécifiques d'une signature afin de juger de l'authenticité de celle-ci.

- Méthode 1 :

Locard ([6]), propose deux classes de caractéristiques pertinentes des signatures manuscrites. La première classe regroupe les caractéristiques de la signature authentique qui sont à la fois peu visibles par le faussaire et difficiles à imiter. Par exemple la qualité du tracé et les alignements. Alors que les caractéristiques de la deuxième classe sont visibles et faciles à imiter. Par exemple, le tracé général des lettres, l'orientation de la signature dans le plan et sa position relative à la marge gauche du texte. Nous décrivons plus précisément trois caractéristiques proposées par Locard ([6]) et utilisées par des méthodes automatiques de vérification de signatures (Huang et al. [13], Mizukami et al. [20], Fang et al. [23]).

Les alignements

Selon Locard, les alignements sont très fiables pour la détection des faux spécimens de signatures manuscrites. On retrouve deux alignements dans une signature, l'alignement inférieur, défini par la ligne de base des lettres et l'alignement supérieur, défini par la ligne enveloppant tous les sommets de la signature. En général, les alignements sont peu variables dans les signatures authentiques. Peu de faussaires prennent en compte ces alignements car ils sont extrêmement difficiles à imiter. Cependant ils ne sont pas identifiables quand il s'agit d'une signature de type européen.

La nature du tracé

La nature du tracé est indépendante de la forme des lettres. Elle est caractérisée par l'alternance de traits ascendants et descendants, ainsi que par les variations conjointes de pression et de la vitesse. Cependant, la nature du tracé est fortement influencée par le choix de l'outil d'écriture. Par exemple, l'utilisation d'un stylo à bille produit des traits réguliers car l'encre est déposée uniformément sur la feuille à l'aide de la bille du stylo, alors que la plume produit un trait irrégulier car l'encre coule directement sur le papier.

Le mouvement

Le mouvement a une influence directe sur la graphie de la signature. Il en résulte une variation dans les proportions[2] de la signature, la présence de tremblements dans le tracé et la présence de liaisons donc l'absence des levées de plumes. Ces caractéristiques sont indépendantes du choix de l'outil de l'écriture.

- Méthode 2 :

Slyter ([7]) propose une méthode basée sur une balance (ratio) entre le rythme et la forme de la signature. En effet, il considère l'écriture, et la signature en particulier, comme une combinaison de plusieurs éléments, les uns élémentaires, les autres plus complexes à identifier et à mesurer, qui interagissent sous l'influence de l'habitude. Il représente cette interaction à l'aide d'un ratio entre les caractéristiques dynamiques et statiques. Pour qu'une fausse signature soit considérée comme vraie il faut que le faussaire imite à la fois le rythme et la forme de l'originale.

Des deux aspects, le rythme est le plus difficile à reproduire et il est le plus représentatif de l'unicité de la signature. Nous énumérons dans ce qui suit les principales mesures décrivant le mouvement associé à une signature.

Vitesse

Il est difficile de retrouver l'information sur la vitesse lorsque nous sommes en présence d'une signature statique. Cependant, certains éléments fournissent une bonne indication sur la vitesse de l'écriture. Ainsi, un mouvement rapide produit des traits clairs ou entre-coupés par endroit, alors qu'un mouvement lent produit un trait plus foncé et continu. Des traits longs qui changent graduellement de niveau d'intensité sont indicateurs d'un changement de vitesse.

La continuité des courbes fournit une autre indication sur la vitesse. Une courbe dessinée avec un mouvement lent comporte des arrêts brusques, alors qu'un mouvement rapide produit des courbes lisses.

Finalement, les changements de pression fournissent une bonne approximation sur la vitesse de l'écriture. La pression appliquée par le scripteur est indiquée par la largeur et la densité des lignes. La majorité des scripteurs augmentent ou diminuent la pression exercée sur le papier dépendamment du sens des traits. En effet, les scripteurs ont tendance à appliquer une forte pression sur un trait montant, alors qu'ils dessinent les traits descendants à main levée. Les zones de pression dans une signature est une caractéristique majeure pour l'authentification et sera discutée dans une prochaine sous-section. Le changement graduel de la pression indique que la vitesse du stylo évolue normalement, alors qu'un changement anormal de la vitesse provoque une rupture dans le rythme de la signature ce qui le rend facilement repérable.

Selon Slyter ([7]) les paramètres reliés à la vitesse (changements de pression, la continuité des courbes...) forment un premier groupe de caractéristiques à considérer lors d'une authentification étant donné l'importance des informations qu'ils fournissent sur le rythme du tracé d'une signature.

Proportions

L'habitude influence l'acte de produire une signature avant même de toucher le papier. L'utilisation de l'espace alloué et l'emplacement par rapport à une ligne fixe[3] font partie de l'habitude du scripteur. L'étude minutieuse d'un ensemble d'exemplaires d'une même signature indique l'endroit où le signataire à l'habitude de placer sa signature par rapport à la ligne fixe. Elle révèle aussi l'influence de l'espace alloué pour la signature sur la largeur de chacune de ses lettres. L'espace blanc entre les différentes parties de la signature est lui aussi une caractéristique de l'habitude du signataire qu'on peut mesurer en le comparant avec la largeur de la signature. Les proportions de la signature (largeur et hauteur) sont très variables et dépendent généralement de l'espace alloué.

Un autre élément important de cette classe est la relation entre les différentes zones de la signature. Slyter ([7]) subdivise la signature en trois zones, la zone supérieure, la zone inférieure et la zone du milieu. Il analyse ensuite, le changement de la hauteur de la zone du centre tout le long de la signature, ainsi que la taille de la zone supérieure par rapport à la taille de la signature. Ces changements sont une conséquence de l'habitude et peuvent être calculé en utilisant l'enveloppe supérieure de la signature.

La ligne de base est une autre caractéristique des proportions de la signature. La ligne de base suit l'inclinaison générale de la signature, l'angle qu'elle forme avec une ligne horizontale[4] est une caractéristique directe de l'habitude du signataire. La majorité des scripteurs ont tendance à écrire avec une inclinaison vers la droite, rares sont les gens qui peuvent produire une écriture avec une inclinaison vers la gauche.

---

[3] La ligne fixe est la ligne qu'on retrouve dans les chèques et les formulaires pour indiquer l'endroit où le scripteur doit signer

Pression

L'intensité de la pression exercée par l'outil d'écriture est contrôlée inconsciemment par le scripteur. Certains signataires appliquent peu de pression sur l'instrument d'écriture, alors que d'autres poussent fortement l'instrument sur le papier. Imiter correctement la pression du stylo reste très difficile. En effet, la pression du stylo varie graduellement lors du tracé de la signature et produit des changements dans la densité et la largeur du trait.

La signature représente un dessin graphique composé d'un ensemble de lettres et de mouvements. La forme des lettres appartenant à une signature diffère d'un mot à un autre. Plus le nombre des lettres connectées dans une même signature est grand, plus il est difficile d'imiter sa forme. Quelques éléments permettent de faire une comparaison de forme entre deux signatures.

Ainsi, le design général des lettres ou des ensembles de lettres est difficile à reproduire surtout si leur longueur est considérable. En outre, la connectivité entre les différentes parties de la signature est également importante. Slyter ([7]), s'intéresse particulièrement à la forme des traits qui sert à connecter chaque lettre avec celle qui la suit et sur leur emplacement.

Finalement, les variations de taille que subissent les lettres dépendamment de leur emplacement dans la signature, sont un bon moyen de comparaison.

2.3 Systèmes de vérification automatique

Dans cette section nous présentons une description des différentes composantes d'un système d'authentification de signatures et nous décrivons brièvement les principaux travaux effectués à ce jour sur l'authentification automatique des signatures manuscrites. Nous préciserons, en regard de ces travaux, les objectifs de notre travail dans ce domaine.

2.3.1 Description d'un système d'authentification de signatures

Un système automatique d'authentification des signatures manuscrites est un système informatique capable de vérifier l'authenticité d'une signature apposée sur un document. Un tel système est généralement constitué de deux modules, soit un module d'acquisition et un module d'authentification. On retrouve à la figure 2.3 un schéma décrivant un tel système à deux modules ainsi que les composantes de chacun des modules.

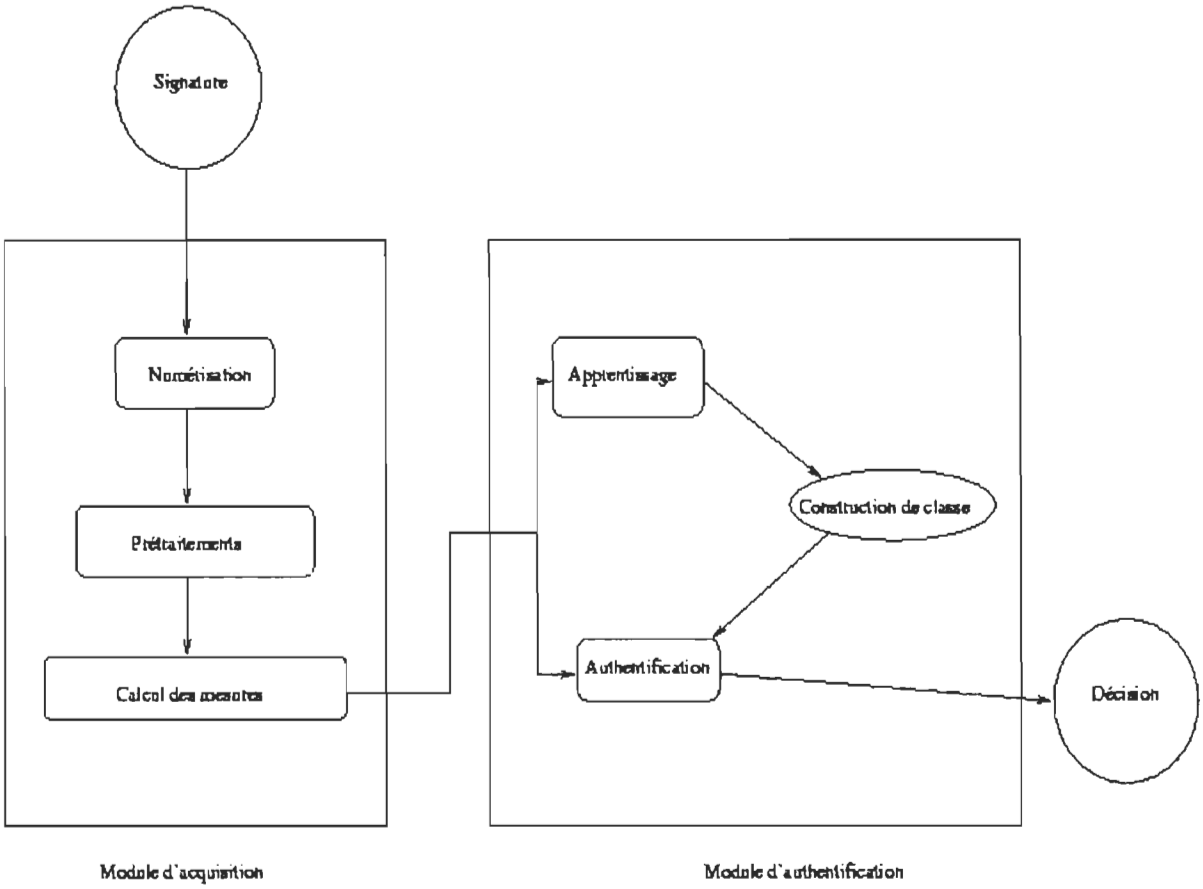

Figure 2.3: Synoptique d'un système d'authentification de signatures.

Une première opération, dans le module d'acquisition, consiste à numériser le document contenant la signature à l'aide d'un scanner, d'une caméra CCD ou d'une barrette CCD. Après cette étape de numérisation, une étape de prétraitement permet d'éliminer le bruit et toute autre information sans lien avec la signature (montants des chèques, lignes de chèques...). Les caractéristiques (mesure des paramètres) utiles au module d'authentification sont alors extraites avant le passage au module d'authentification.

Dans le module d'authentification on distingue deux sous modules : soit une phase d'apprentissage et une phase d'exploitation. Nous expliquons dans le paragraphe suivant le déroulement de la phase d'apprentissage pour deux types de systèmes largement utilisés dans la littérature (R.Plamondon et al.[27]) : soit les systèmes basés sur une classification statistique et ceux basés sur une classification par réseau de neurones.

Phase d'apprentissage.

La phase d'apprentissage permet au système de s'adapter progressivement aux caractéristiques propres à un scripteur en analysant un nombre limité de ses signatures qui seront appelées des références.

Deux méthodes sont utilisés pour la mise en œuvre de cette phase (Bajaj et al.[25]). Une première approche est basée sur une statistique tandis que la seconde utilise les réseaux de neurones.

Dans le cas des systèmes basés sur une approche statistique, un représentant pour chaque signataire est construit. Ce représentant peut être constitué d'un ensemble de signatures ou encore d'une «signature moyenne» calculée à partir des signatures de référence. Nous trouvons un exemple de cette approche dans Elyassa ([21]) ou la moyenne et l'écart type statistique des distances prétopologiques entre deux signatures appartenant à l'ensemble de référence, sont calculés.

Dans le cas des méthodes neurales, des signatures sont introduites dans le réseau de neurones. Les poids du réseau sont modifiés en fonction de leurs réponses. Le représentant est donc défini à partir des poids du réseau. Un des intérêts de cette méthode est la réduction de la taille des données relatives à chaque signataire en fonction du nombre de références utilisées.

Phase d'exploitation.

Pour les systèmes basés sur les méthodes statistiques, il s'agit, pendant la phase d'exploitation, de comparer la signature présentée avec la référence du signataire supposé et de proposer une décision d'acceptation ou de rejet de la signature en fonction de cette comparaison.

Dans le cas des réseaux neuronaux, on soumet la signature à vérifier et le système propose une décision en fonction de la réponse du réseau neuronal.

2.3.2 Évaluation d'un système d'authentification

Afin d'évaluer les performances d'un système d'authentification, deux paramètres sont généralement utilisés: le taux de vrais rejetés (TVR) et le taux de faux acceptés (TFA). Ces deux taux sont aussi appelés de type I ou FRR (False Rejection Rate) pour le TVR et type II ou FAR (False Acceptance Rate) pour le TFA (R.Plamondon et al.[27]).

On peut considérer le problème de l'authentification de signatures comme un problème de partitionnement en deux classes. Pour un signataire donné, une première classe est formée par les signatures du signataire et l'autre classe par toutes les autres signatures disponibles. Idéalement, ces deux classes sont séparables par une hypersurface dans l'espace des représentations des signatures (ensemble des paramètres ou mesures).

Le problème de l'authentification se résume essentiellement à trouver la forme et la position de cette hypersurface dans l'espace adéquat. Généralement, les systèmes d'authentification se servent d'un paramètre, appelé seuil de décision, qui permet d'ajuster l'hypersurface. Ainsi, si les deux classes sont séparables, il existe une valeur du seuil de décision qui annule les deux taux d'erreurs TVR et TFA (figure 2.4).

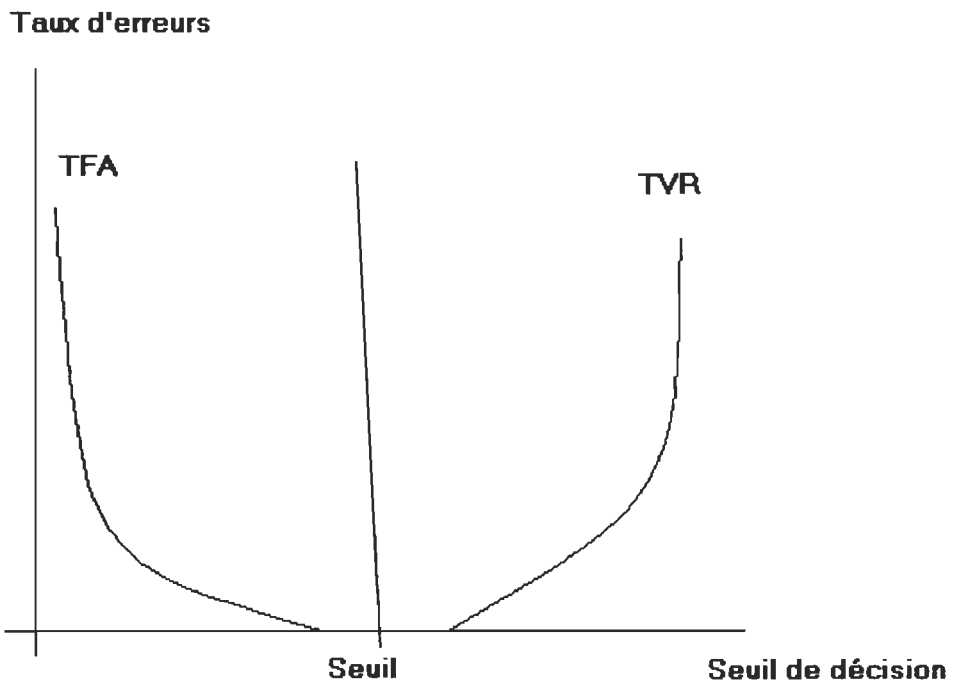

Figure 2.4 : Choix d'un seuil de décision qui annule les taux d'erreurs.

En pratique, la plupart des signatures de références utilisées ne permettent pas la séparation en deux classes disjointes (figure 2.5)

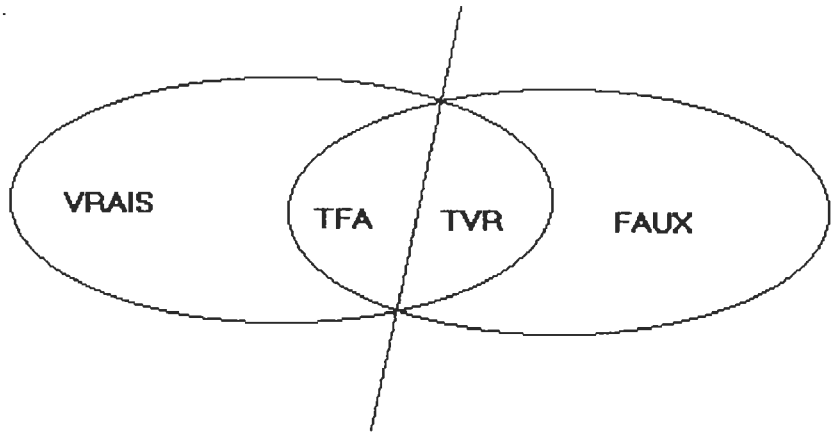

Figure 2.5 : Classes non séparables.

Le choix du seuil de décision est construit à partir d'un des critères suivants (figure 2.6) :

- Minimiser les moyennes des taux TFA et TVR.
- S'assurer que l'un des deux taux soit inférieur à un certain seuil désiré[5] (par exemple inférieur à 1%).

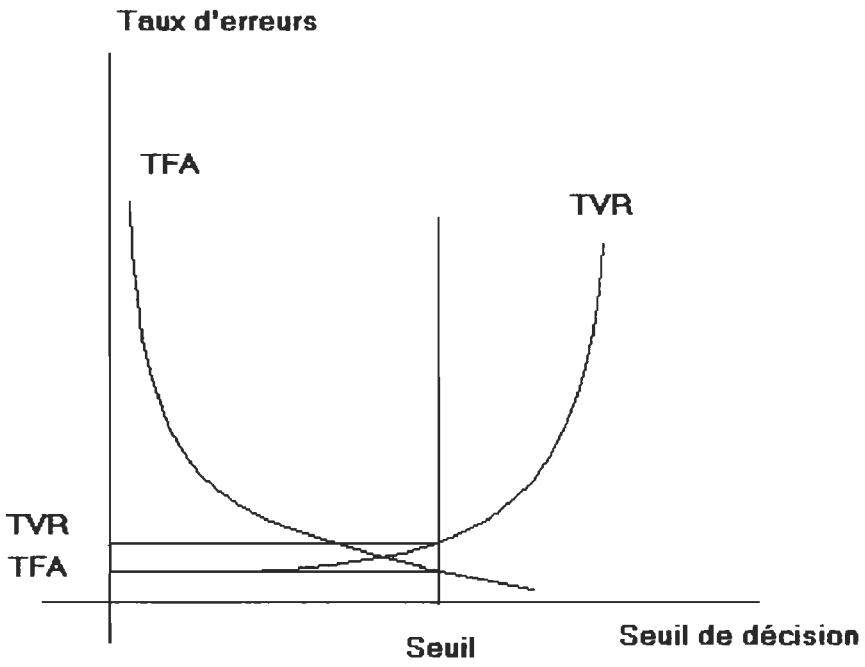

Figure 2.6 : Choix d'un seuil de décision suivant un critère.

[5] Comme le montre la figure 2.6 il est impossible qu' un des deux taux d'erreur soit proche de 0 sans

2.3.3 Méthodes Automatiques

Plusieurs méthodes ont été proposées à ce jour pour développer un système le plus complet possible pour la vérification automatique de la signature manuscrite.

Certains auteurs traitent seulement de la première étape de ce processus, à savoir l'extraction de l'image de la signature à partir d'un chèque bancaire ou d'un document. Par exemple la méthode proposée par Djeziri et al. ([32]) utilise un critère de filiformité afin d'extraire la signature à partir d'une image de chèque. Ce type de problème, étant considéré comme un problème d'extraction, et non, de vérification de signature, ne fait pas partie du présent mémoire. Nous travaillerons donc, avec des images ne contenant que les signatures apposées sur un fond blanc uni (*c.f* Chapitre 4 pour la description détaillée de la base de données).

Afin d'évaluer ultérieurement les performances de notre système, nous ne discuterons dans cette partie que des principaux travaux effectués sur l'authentification automatique des signatures manuscrites. Une classe de méthodes (Burr [8],Sabourin et al.[9], Hunt et al.[15], Ramesh et al.[24], Bajaj et al.[25], Plamondon et al.[27]) utilisent des mesures globales pour caractériser la signature, soit des paramètres géométriques comme :

- les dimensions de la signature : hauteur et largeur de la signature,
- les projections : projection horizontale et verticale,
- la pente : l'inclinaison de la signature par rapport à l'axe horizontale,
- l'enveloppe : contour convexe de la signature.

Généralement, les méthodes utilisant des mesures globales produisent de bons résultats pour vérifier des faux aléatoires, comme par exemple Ammar et al.([10],[11]) qui obtiennent des performances de l'ordre de 10.25% d'erreur moyenne sur une base de données de 200 spécimens. Cependant la performance laisse à désirer lorsqu'il s'agit de faux simples ou par imitation.

Une solution à ce problème serait de considérer des mesures locales pour caractériser les signatures (Wilkinson et al.[28], Sabourin et al.[14]). En effet, Huang et al. ([13]) proposent des mesures locales basées sur l'évaluation de la distribution de l'encre à différentes résolutions dans des régions de forte pression. Pour une signature donnée, le squelette, le contour ainsi que les frontières suivant les directions (Nord, sud, est, ouest) sont calculées et forment un vecteur dont la taille varie suivant la résolution voulue.

Sabourin et al. ([14]), proposent une méthode d'extraction de paramètres locaux à l'aide d'une grille appliquée à l'image. Une base de données de 800 signatures authentiques appartenant à 20 personnes a été construite à cet effet par Sabourin et al. ([14]). Chaque individu signait dans un rectangle de dimensions fixes, en utilisant le même type de stylo et de papier, et ce dans un délai de deux semaines à raison de trois à quatre signatures par jour. Dans ce cas les signatures composant la base étaient de type américain. Chaque signature est centrée dans une image de 512x128 pixels, une grille de rectangles est alors appliquée sur l'image. Une distribution granulométrique est utilisée comme paramètre local pour décrire chaque portion de la signature à l'intérieur d'un rectangle de la grille.

Il faut noter que même si la plupart des systèmes proposés jusqu'ici, présentent de bonnes performances sur les mêmes types de faux, celles-ci diminuent considérablement si toutes les catégories de faux sont utilisées simultanément. La raison principale de cette lacune réside dans la difficulté de trouver un ensemble de mesures adéquat pour représenter tous les types de contrefaçon. Ainsi l'utilisation de mesures globales nous permet d'isoler, tel que souhaité les contrefaçons aléatoires, mais ne nous permet pas de distinguer entre une signature véritable et une contrefaçon habile qui est semblable au niveau de la forme à l'original. D'un autre coté, les systèmes utilisant des mesures locales nous permettent de détecter les faux par imitation mais risquent de classer comme faux des signatures authentiques si celles-ci présentent de grandes variations dans la forme.

Une solution à ce problème est de considérer un système Multi-Expert (MES), ces systèmes sont faits à partir d'une variété d'experts simples, chacun apte à résoudre un problème particulier. Ces systèmes présupposent qu'en combinant les résultats de plusieurs systèmes séparés, il est possible de contourner les lacunes rencontrées par certains systèmes par rapport à certains types de faux.

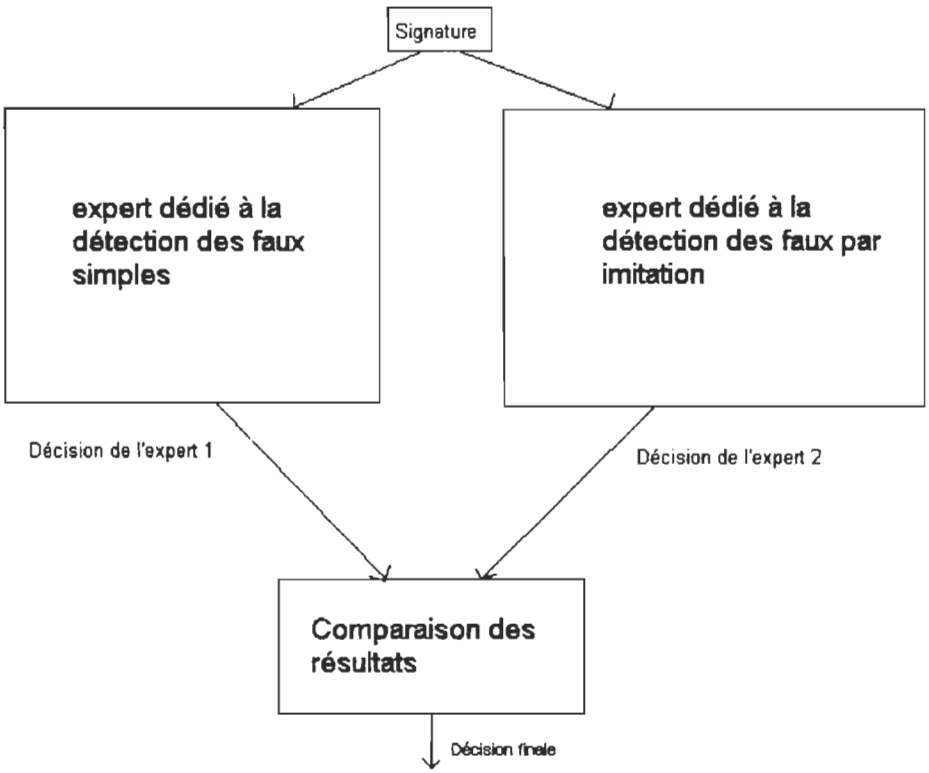

Figure 2.7 : Exemple d'un système MSE.

L'idée d'utiliser les systèmes MSE est récente dans le domaine de la vérification des signatures. La plupart des systèmes existants raffinent leur procédé de décision en adoptant les MSE. Pour ce faire, ils utilisent des mesures globales et locales qu'ils appliquent à l'image à différentes résolutions (Huang et al.[13],QI et al.[15], QI et al.[16]).

C.Sansone et al. ([17]) proposent une méthode basée sur les MSE où chaque expert simple utilise une mesure différente pour caractériser la signature. Le système proposé contient trois experts différents et indépendants les uns des autres. Un premier expert est dédié à la détection des faux simples et un deuxième pour la détection des faux par imitation alors que le troisième combine les résultats des deux systèmes pour donner une décision finale. Les mesures utilisées sont les mêmes proposées par Huang et al. dans ([13]). Baltzakis et al.([19]) utilisent un système MSE à deux experts basé sur l'extraction de plusieurs caractéristiques : mesures géométriques globales (Hauteur, largeur, projections horizontales et verticales, angle d'inclinaison...), en plus d'une mesure basée sur les informations fournies à partir d'une grille appliquée sur la signature ainsi qu'une mesure sur la texture de l'image. Le premier expert du système comporte 4 réseaux de neurones trois pour chacune des trois caractéristiques (mesures globales, grille, texture) et un qui traite la distance euclidienne des trois caractéristiques. Le résultat du premier expert est transmit au deuxième qui utilise un réseau de type RBF (radial basis function) pour fournir une décision finale sur l'authenticité de la signature.

Toutes les méthodes citées plus-haut proposent des mesures géométriques globales ou locales pour caractériser les signatures. Cependant une nouvelle famille de méthodes caractérise la signature par différentes mesures qui tiennent compte de ses variations.

Ainsi, une classe de méthodes utilise la représentation par ondelette pour caractériser les signatures à différentes résolutions (Fadhel et al.[18], McMormack et al.[26]). Fadhel et al. ([18]) utilisent la transformée par ondelettes du squelette de la signature qui permet d'obtenir une interprétation unique de la signature à différentes résolutions, basée sur des coefficients matriciels insensibles à la rotation et au déplacement de la signature.

La méthode de Mizukami et al. ([20]) repose sur l'extraction d'une fonction de déplacement calculée en comparant une signature erronée avec une signature authentique. Le système calcule deux mesures de dissimilitude : la première entre la signature à vérifier et une signature authentique, et la deuxième entre deux signatures authentiques puis compare les deux mesures.

Fang et al. ([23]) introduisent une nouvelle approche qui utilise une fonction de régularité du trait pour authentifier les signatures. Elle est inspirée de l'approche des experts pour authentifier les signatures et utilise un critère de régularité du trait. La signature est divisée en segments continus, et sur chaque segment on considère deux fonctions correspondant aux coordonnées horizontale et verticale. Pour chacune de ses fonctions est calculée une fonction "spline" qui est comparée à sa version originale et le nombre de points d'intersection entre les deux est calculé.

Des mesures basées sur la théorie mathématique des ensembles appliquée au cas discret ont aussi été développées pour caractériser les signatures. Par exemple, Sabourin et al. ([14]) utilisent la morphologie mathématique. Elyassa et al. ([21]) proposent une approche originale basée sur un formalisme prétopologique qui intègre un critère d'agrégation élaboré à partir d'une pseudo-distance. Une distance prétopologique entre deux signatures est calculée, et à partir de celle-ci une décision est prise sur l'authenticité d'une signature. Elyassa et al. ([21]) ont divisé une base de données de 800 signatures en deux collections P_1 et P_2. La collection P_1 est utilisée dans la phase d'apprentissage alors que la deuxième est utilisée dans la phase d'expérimentation. Un critère de décision est défini pour chaque signataire j à l'aide de la pseudo-distance introduite par Lamur ([22]). Pour chaque signataire j, un ensemble de p seuils $(\alpha^1,...,\alpha^p)$ est choisi et une classe C_j est créée, où $C_j=(S_j^1,...,S_j^p)$ représente un ensemble de p signatures de l'individu j. Pour une signature S quelconque le critère de décision s'écrit alors :

$$\left(\forall k = 1,..,p;\, d(S_j^k, S) > \alpha^k\right) \Rightarrow S \notin C_j,$$

où d est la pseudo-distance définie par Lamur ([22]). Les représentants et les seuils de chaque classe sont choisis et calculés par apprentissage en utilisant la classe P_1. Pour ce faire, les distances intra-classes, la distance moyenne $\overline{d_j}$ et l'écart type σ_j sont calculés. Les p plus petites valeurs des distances moyennes déterminent les signatures à retenir comme des représentants de la classe C_j. Les seuils α^k $(1 \le k \le p)$ sont posés comme suit;

$$\alpha^k = \overline{d_k} + \mu \times \sigma_k .$$

L'entier p et la valeur μ sont choisis par expérimentation de manière à minimiser les taux de faux acceptés (TFA) et de vrais rejetés (TVR).

Les performances d'un système de vérification de signatures manuscrites dépendent en grande partie des mesures utilisées pour caractériser la signature. Mais, pour avoir un système efficace il est important d'avoir un module de classification performant.

Certains méthodes proposent une classification de type probabiliste qui repose sur la quantification du vecteur des mesures avec des méthodes statistiques (Elyassa et al.[21]). Ces méthodes sont efficaces mais exigent le stockage de plusieurs échantillons de signatures de référence.

Dans d'autres méthodes, les classes de signatures sont identifiées à l'aide de réseaux de neurones (Burr [8], Huang et al. [13], Baltzakis et al. [19], Bajaj et al. [25], Lee et al. [26], Cardot [33]). Ainsi Bajaj et al. ([25]) utilisent une classification basée sur un filet neuronal alimenter vers l'avant (forward), alors qu'un "perceptron" à deux étages est utilisé par Baltzakis et al. ([19]). Huang et al. ([13]) utilisent un perceptron multicouche. Ces méthodes produisent des systèmes intelligents capables de s'adapter automatiquement au changement des habitudes du signataire (Bajaj et al. [25], Lee et al. [26], Cardot [33]), en plus de réduire la taille des échantillons à stocker.

Une classification par des algorithmes génétiques a été proposée par V.E. Ramesh et al. ([24]). Ils utilisent une base de données de 650 signatures produites par 15 individus sur une période d'un mois. Le système calcule 4 types de caractéristiques, chacune traitée par un sous-système indépendant. Une combinaison des résultats des 4 sous-systèmes fournit un résultat final sur l'authenticité de la signature. Les caractéristiques utilisées sont de types : géométriques (ration hauteur/largeur, angle d'inclinaison, projection horizontal...), représentation des moments (Bajaj et al.[25]), enveloppes (Bajaj et al.[25]) et finalement une représentation par coefficient d'ondelette de la signature. Cette méthode présente un taux de faux acceptés (TFA) de 15% et un taux de vrais rejetés (TVR) de 17%.

Dans toutes les approches proposées, la base de données constitue un élément majeur dans l'apprentissage et la vérification du système d'authentification. Pour valider correctement un système d'authentification, il est nécessaire de disposer d'une base de données qui représente toute la communauté avec ses classes sociales (Plamomdon et al.[27], Evette et al.[36], Evette et al.[37]) démontrent que chaque signataire change sensiblement sa signature avec le temps, il faut donc que la base des signatures de référence soit continuellement mise à jour pour assurer le bon fonctionnement des systèmes d'authentification.

On retrouve ci-après un tableau comparatif des performances et des caractéristiques des méthodes citées dans cette section du mémoire.

Auteurs, Année	Méthodes	Base de Données	Résultats
K. Huang et al. 1997	- Mesures géométriques; - squelette, - contour, - frontières. - Réseaux de neurones pour la classification.	- 21 individus. - 24 authentiques. - 144 faux. - Total 3528.	TVR: 11,1% TFA: 11,8%
Sabourin et al. 1997	- Distributions granulométriques. - Classificateur à seuil. - Classificateur utilisant la méthode du plus proche voisin.	- 20 individus. - 40 authentiques. - Total 800.	- 0.02% d'erreur pour le plus proche voisin - 1.0% d'erreur pour le seuillage
Fadhel et al. 1999	- Calcul les coefficients par ondelettes du squelette de la signature. - Utilise un réseau de neurones multicouches pour la vérification.	- 30 individus. -10 authentiques. - Total 300.	TVR: 6.2% TFA: 5.5%

Auteurs, Année	Méthodes	Base de Données	Résultats
Y.Mizukami et al. 1999	- Calcul de deux mesures de dissimilitude : la première entre la signature à vérifier et une signature authentique, et la deuxième entre deux signatures authentiques - Comparaisons des deux mesures.	- 20 individus. -200 authentiques. -200 faux simple. -Total 400.	Erreur moyenne de 24,9%
B.Fang et al. 1999	La méthode est inspirée de l'approche des experts pour authentifier les signatures et utilise un critère de régularité du trait. Une fonction "spline" est calculée puis comparée à sa version régulière. Le nombre de points d'intersection entre les deux permet de décider de l'authenticité d'une signature.	-55 individus. - 24 authentiques. - 24 faux par imitation. - Total 2640.	Erreur moyenne de 21.7%
V.E Ramesh et al. 1999	Caractéristiques : - géométriques, - représentation des moments, - enveloppes, - une représentation en ondelette. Classification : algorithmes génétiques.	- 15 individus. - 43 authentiques. - Total 650.	TVR=17% TFA=15%
C.Sansone et al. 2000	Trois systèmes différents : - un système dédié à la détection des faux simples, - un système pour la détection des faux par imitation, - un troisième combine les résultats des deux systèmes pour donner une décision finale. Les mesures utilisées sont les mêmes proposées dans (Huang et al.[13]).	- 49 individus. - 40 authentiques. - Total 1960.	TVR: 5.71% TFA: 4.29%

Auteurs, Année	Méthodes	Base de Données	Résultats
H.Baltzakis et al. 2001	Caractéristiques : - mesures géométriques, - grille appliquée sur signature, - texture. Classification : - réseau de neurones à deux étages.	- 115 individus. - 15 à 20 authentiques. - Total 2000.	Erreur moyenne de 19%
M.Elyassa et al. 2001	- Calcul une distance prétopologique pour décider de l'authenticité d'une signature. - Calcul la moyenne et l'écart type des signatures obtenues pour la classification.	- 20 individus. - 40 authentiques. - Total 800.	TFA: 1,75% TVR: 4,00%.

Chapitre 3

Pré-traitement des signatures

Dans ce chapitre nous présentons de façon détaillée les différentes étapes de prétraitement utilisées. Ces étapes consistent à filtrer les données qui seront utilisées par le système d'authentification mis au point et présenté au chapitre 4. Dans une première partie nous décrivons les bases de données utilisées lors de cette étude. Dans la seconde partie nous présentons les définitions relatives au concept de morphologie mathématique et finalement la troisième partie portera sur les opérations de pré-traitements appliquées aux images de signatures.

3.1 Bases de données

Nous avons utilisé deux bases de données distinctes pour la mise au point et la validation de la méthodologie de reconnaissance présentée dans ce mémoire. La première banque de signatures a été construite par Sabourin et al [14]. Cette base contient 800 signatures, soit 40 spécimens de 20 signataires. Chaque signature est numérisée en niveau de gris et de dimensions 512 x 128 pixels. Cette base est scindée en deux sous-ensembles P_1 et P_2 de 400 signatures chacune. Les éléments de P_1 sont choisis aléatoirement et sont utilisés pour la phase d'apprentissage du système d'authentification. Le sous-ensemble P_2 est utilisé lors de la phase d'expérimentation. Toutes les signatures contenues dans cette base sont du type américain.

La seconde base de donnée a été construite par l'équipe de F.Nouboud (UQTR) et elle contient des signatures de type américain et européen. Elle est constituée de 2600 signatures, produites par 130 signataires à raison de 20 signatures par personne. Chaque image est en niveau de gris et de dimensions variables.

3.2 Rappels de morphologie mathématique

Dans cette section, nous rappelons les définitions et propriétés des opérateurs fondamentaux de la morphologie mathématique que nous introduisons dans un cadre ensembliste adapté aux images binaires.

Dans ce qui suit, nous utiliserons les notations suivantes;

- $P(R^n)$: l'ensemble des parties de l'espace euclidien R^n,
- X^c : le complémentaire dans R^n d'un ensemble $X \subset P(R^n)$,
- X' : le transposé de l'ensemble $X \subset P(R^n)$, définit par : $X' = \{-x,\ x \in X\}$.

3.2.1 Extensivité - Monotonie - Idempotence

Soit une transformation $\psi : P(R^n) \to P(R^n)$.

- On dira que ψ est *extensive* si et seulement si :

$$\forall X \in P(R^n),\ X \subseteq \psi(X).$$

- On dira que, ψ est *anti–extensive* si et seulement si :

$$\forall X \in P(R^n),\ \psi(X) \subseteq X.$$

- La transformation ψ est dite *croissante* si et seulement si elle préserve la relation d'ordre naturelle sur $P(R^n)$ (inclusion) :

$$\forall (X,Y) \in P(R^n)^2,\ X \subseteq Y \Rightarrow \psi(X) \subseteq \psi(Y).$$

- On dira que ψ est *idempotente* si et seulement si :

$$\exists p > 0,\ \forall X \in P(R^n),\ \psi \circ \psi^p(X) = \psi^p(X).$$

- On appellera *filtre morphologique* toute transformation croissante et idempotente.

- Enfin, deux transformations ψ_1 et ψ_2 seront dites *duales* si et seulement si :

$$\forall X \in P(R^n),\ \psi_1(X) = (\psi_2(X^c))^c.$$

3.2.2 Érosion et dilatation

- On appelle le *dilaté* d'un ensemble X par l'élément structurant[6] B l'ensemble noté $\delta_B(X)$ et défini par :

$$\delta_B(X) = X \oplus B^{\vee} = \{x \in R^n, X \cap B_x \neq \phi\},$$

où $\oplus$ désigne l'adition de Minkowski.

- On appelle l'*érodé* de X par l'élément structurant B l'ensemble noté $\varepsilon_B(X)$ et définit par :

$$\varepsilon_B(X) = X \Theta B^{\vee} = \{x \in R^n, B_x \subset X\},$$

où Θ désigne l'adition de Minkowski.

Les opérateurs $\delta_B(X)$ et $\varepsilon_B(X)$ sont respectivement appelés dilatation et érosion par rapport à l'élément structurant B.

3.2.3 Ouverture et fermeture

Les opérateurs d'*ouverture* et de *fermeture* sont des transformations morphologiques de $P(R^n)$ dans $P(R^n)$ obtenus par composition des opérateurs élémentaires d'érosion et de dilatation.

Par définition, l'*ouverture* d'un ensemble relativement à un élément structurant B est la transformation, notée γ_B, obtenue par composition d'une érosion suivie d'une dilatation :

$$\gamma_B = \delta_B \circ \varepsilon_B .$$

On montre aisément que cette transformation est *anti-extensive.*

La *fermeture* relativement à un élément structurant B est transformation, notée ϕ_B ,obtenue par la composition d'une dilatation suivie d'une érosion :

$$\phi_B = \varepsilon_B \circ \delta_B .$$

Cette transformation est *extensive.*

3.3 Pré-traitement

Cette étape est importante en reconnaissance des formes et donc lors du processus d'authentification des signatures. Cette étape consiste à préparer les données qui seront fournies au module d'authentification. La qualité de ces données influe beaucoup sur les résultats finaux.

Lors de la saisie des signatures à l'aide d'un scanner ou d'un appareil photo numérique, du bruit peut être introduit et nuire à la qualité de l'image obtenue et donc à son interprétation. La structure même d'une signature peut être modifiée, par exemple, par l'élimination de certaines parties de la signature ou encore par l'ajout d'artéfacts qui seront incorporés à la signature.

Plusieurs traitements sont nécessaires pour passer de l'image bruitée à une image améliorée de la signature. Nous présentons dans ce paragraphe les différents filtres utilisés lors de notre étude.

3.3.1 Binarisation

Après l'acquisition numérique d'une signature, nous disposons d'une image à 256 niveaux de gris. Pour comparer deux signatures, nous nous intéressons principalement à la forme du tracé, et donc nous n'avons besoin que de deux couleurs seulement: une pour représenter le fond (blanc) et une autre pour le tracé (noir). Une première étape consiste donc à binariser l'image de la signature. Dans le cas où nous souhaiterions extraire des caractéristiques pseudo-dynamiques[7] de la signature, il faudrait alors, travailler avec une image en niveaux de gris.

Cette opération de binarisation n'est pas facile car elle dépend des nuances dans le tracé, de l'encre utilisé, du fond, de l'éclairage lors de l'acquisition, etc.

Il faut donc effectuer un seuillage sur les niveaux de gris afin de partager les niveaux de gris en deux classes, soit celle du fond et celle du tracé. Un mauvais choix du seuil entraînera un bruit important ou une perte d'une partie du tracé. Ce seuil peut être variable pour chaque image ou encore pour les différentes parties d'une image, cela afin de tenir compte de la non-uniformité de l'éclairage lors de la capture de l'image.

[7] Plusieurs méthodes proposent des systèmes de vérification de signature Off-Line qui tentent de retrouver et d'utiliser des informations dynamiques tels que la vitesse et la pression. Ces caractéristiques sont

Nouboud ([40]), après une étude en fonction du type de stylo, de la couleur de l'encre et du niveau de gris d'un fond uniforme, propose comme seuil de binarisation, $S = N_f - 2$, où N_f est le niveau de gris du fond. Nous obtenons ainsi, quand le fond est uniforme, une binarisation satisfaisante pour la plupart des images.

Cette méthode est efficace pour les signatures apposées sur une feuille blanche, mais dans le cas des images avec un fond non uniforme, il est préférable d'utiliser un seuillage adaptatif. A cet effet un seuil est calculé à partir de l'histogramme des niveaux de gris de l'image. Nous avons retenus cette approche dans notre étude.

La méthode proposée à été développée par Nouboud et al ([41]) et tire partie d'une part de l'histogramme de l'image et d'autre part, du voisinage en 8-connexité du point à binariser. Le seuil de binarisation est un seuil variable. Il est calculé pour chacun des points à traiter, en prenant en compte la configuration du voisinage.

Cette binarisation s'effectue en deux étapes. La première étape consiste à établir un seuil par défaut, par la suite en modifiant ce dernier les seuils de binarisation sont calculés pour chaque pixel. Ce seuil par défaut est défini à partir de l'histogramme de l'image de la signature. À partir de la configuration du voisinage du pixel à traiter nous déduisant la valeur d'un biais à ajouter au seuil.

Pour déterminer la valeur à ajouter au seuil par défaut pour chaque pixel nous effectuons la binarisation d'un voisinage 3x3 du pixel étudié à l'aide du seuil par défaut, puis le nombre de pixels noir issus de cette première binarisation et présents dans le voisinage est calculé. En se basant sur ce nombre de pixels noirs, nous calculons le biais à ajouter au seuil par défaut en utilisant les relations suivantes :

$$S_0 = S + \frac{255 - S}{2} \qquad S_5 = S - \frac{S}{2}$$

$$S_1 = S + \frac{255 - S}{4} \qquad S_6 = S - \frac{S}{4}$$

$$S_2 = S + \frac{255 - S}{8} \qquad S_7 = S - \frac{S}{8}$$

$$S_3 = S + \frac{255 - S}{16} \qquad S_8 = S - \frac{S}{16}$$

$$S_4 = S$$

où S_i est le seuil du pixel central si son voisinage contient i points noirs et S est le seuil défini par défaut. Pour terminer, nous effectuons la binarisation du point considéré à l'aide de ce seuil biaisé. Les figures suivantes montre le résultat de la binarisation effectué à l'aide des deux méthodes.

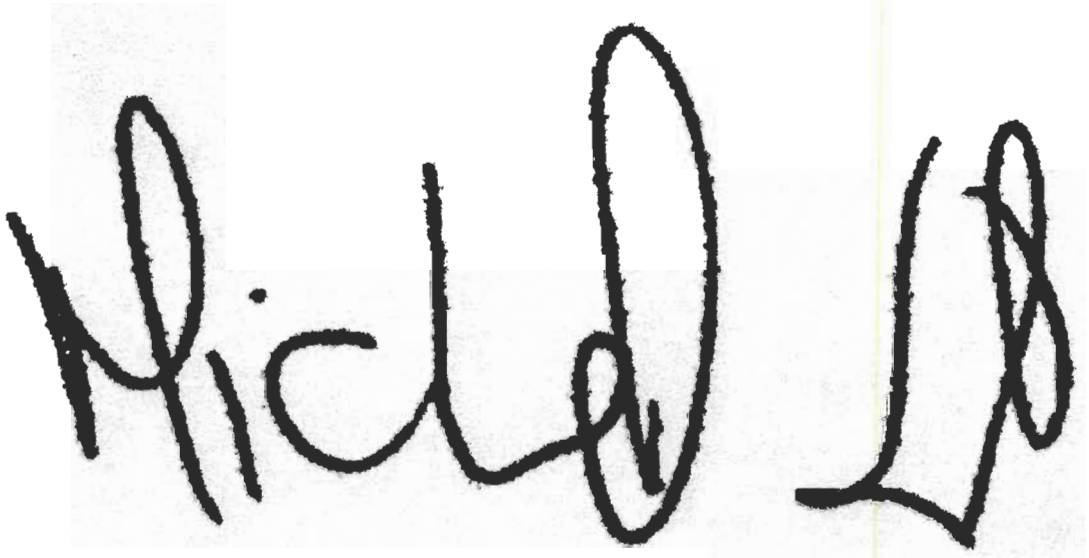

Figure 3.1 : Image originale.

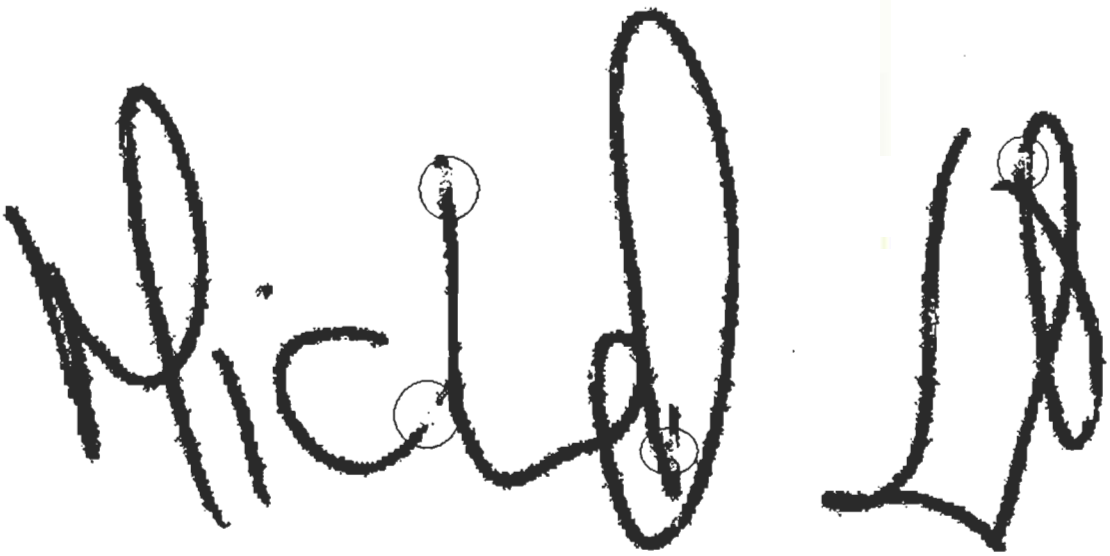

Figure 3.2 : Binarisation brut.

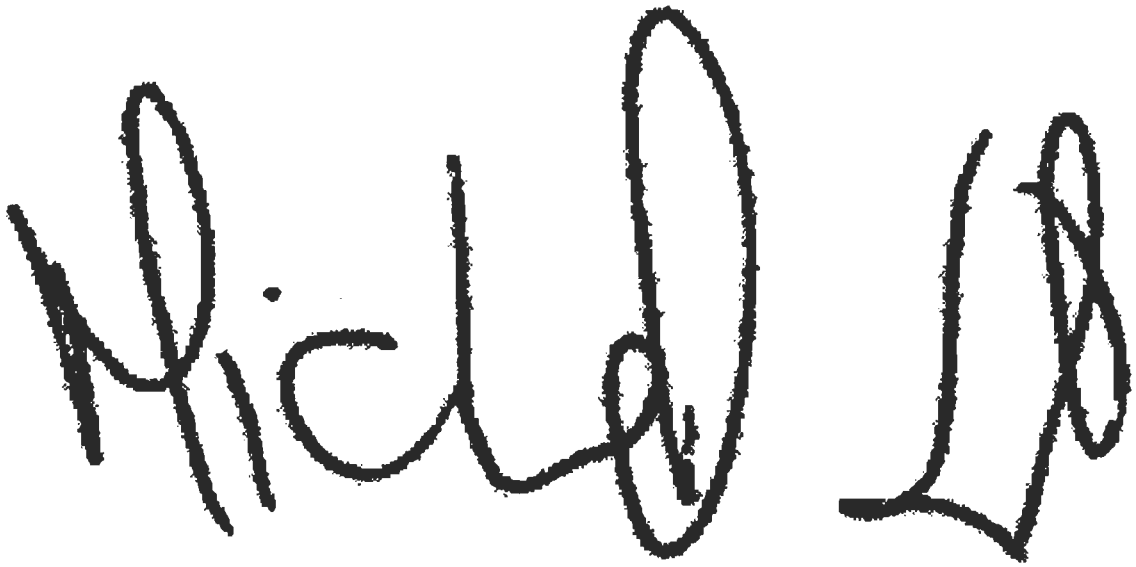

Figure 3.3 : Binarisation avec seuil variable.

La figure 3.2 représente une signature binarisée à l'aide de la première méthode, alors que la figure 3.3 montre la même signature binarisée à l'aide de la deuxième méthode. Les cercles sur la première image indiquent les régions de la signature affectées par le changement de méthode de binarisation.

3.3.2 Élimination du bruit

Après l'étape de la binarisation, il se peut que des points n'appartenant pas à la signature soient présents. Afin d'éliminer ces points nous appliquons des filtres morphologiques de base. En l'occurrence nous appliquons sur l'image binarisée un opérateur d'ouverture de taille 2 suivi d'un opérateur de fermeture de taille 2. Ces filtres ont pour objectif l'élimination des points noirs n'appartenant pas à la signature (*élimination du bruit*) et de remplir les points blancs qui apparaissent dans le trait de la signature (*remplissage de trous*) sans affecter la structure du tracé de la signature.

3.3.3 Définition de la fenêtre de travail

Puisqu'une signature n'occupe généralement pas toute l'image numérisée, il est important de définir une fenêtre de travail englobant la signature, sur laquelle l'ensemble des traitements sera effectué. Cette standardisation réduira considérablement le temps de traitement par le système d'authentification.

Afin de définir cette fenêtre nous recherchons dans l'image les bornes transversales de la signature (*voir figure 3.4*). Nous obtenons ces bornes en parcourant horizontalement et verticalement l'image à la recherche du premier pixel noir; la position de ce pixel représente un des bords de la fenêtre.

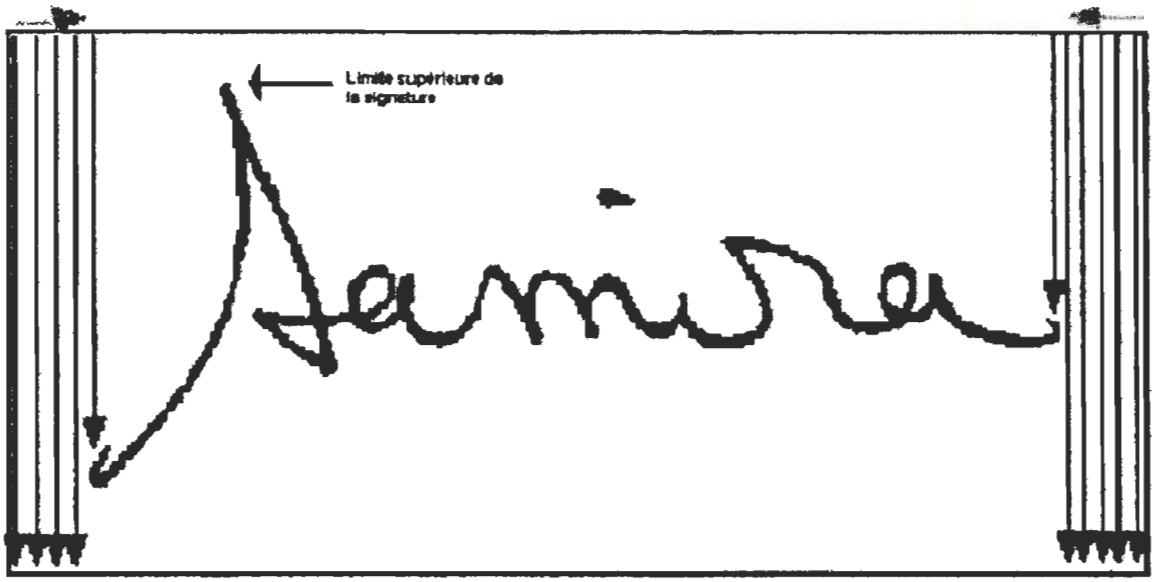

Figure 3. 4 : Recherche des limites d'une signature.

Nous obtenons à la fin de cette opération un rectangle qui encadre la signature et qui passe par ces limites.

Chapitre 4

Méthode d'authentification automatique

Dans ce chapitre nous présentons deux systèmes d'authentification automatique qui ont été implémentés et testés. Un premier système est dédié à la reconnaissance du type de la signature (américaine ou européenne) alors que le deuxième est une amélioration d'une méthode de vérification de signatures basée sur la pré-topologie mathématique et qui fut introduite par Elyassa et al ([21]).

4.1 Reconnaissance du type de la signature

Notre premier système reconnaît le type des signatures. Tel que cité dans le paragraphe 2.1.2, les signatures américaines s'apparentent à l'écriture cursive, alors que les signatures européennes possèdent une composante graphique importante. Pour cela nous utilisons des mesures géométriques largement décrites dans le domaine de la vérification automatique des signatures manuscrites (Ammar [11], Bajaj et al.[25], Plamondon et al.[27], Doria et al [43]) pour leur efficacité et pour la facilité de leur implémentation. Les mesures utilisées sont :

- les moments,
- les projections,
- le ratio hauteur-largeur,
- le nombre de parties de la signature.

Le choix des moments invariants (HU [42], Doria et al [43]) comme mesure est dû principalement à leurs propriétés de stabilité par rapport aux distorsions des dimensions de la signature et au changement de position et de direction.

L'équation générale du moment bi-dimensionnel m_{pq} pour une fonction $f(x,y)$ est :

$$m_{pq} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^p y^q f(x,y)\,dx\,dy \,. \quad (4.1)$$

Dans le cas discret, pour une image $g(x,y)$ de dimension (N x M), cette équation s'écrit :

$$m_{pq} = \sum_{y=0}^{M-1}\sum_{x=0}^{N-1} x^p y^q g(x,y). \quad (4.2)$$

Hu ([42]) a introduit six moments invariants par rapport à la rotation et à la translation (Invariant Moment) définis comme suit :

$$\begin{aligned}
M_1 &= m_{20} + m_{02}\,,\\
M_2 &= (m_{20} - m_{02})^2 + (3\,m_{21} - m_{03})^2\,,\\
M_3 &= (m_{30} - 3\,m_{12})^2 + (3\,m_{21} - m_{03})^2\,,\\
M_4 &= (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2\,,\\
M_5 &= (m_{30} + 3\,m_{12}) + (m_{30} + m_{12})\left[(m_{30} + m_{12})^2 - 3\,(m_{21} + m_{03})^2\right] +\\
&\quad (3\,m_{21} - m_{03})^2\,(m_{21} + m_{03})\left[3\,(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2\right],\\
M_6 &= (m_{20} - m_{02})\left[(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2\right] + 4\,m_{11}\,(m_{30} + m_{12})(m_{21} + m_{03})\,,
\end{aligned} \quad (4.3)$$

où m_{ij} est défini en (4.2). Les moments que nous avons utilisées sont ceux présentées par Doria et al ([43]) :

$$\beta_1 = \frac{\sqrt{M_2}}{M_1}\,, \qquad \beta_2 = \frac{M_3 m_{00}}{M_1 M_2}\,, \qquad \beta_3 = \frac{M_4}{M_3}\,,$$

$$\beta_4 = \frac{\sqrt{M_5}}{M_4}\,, \qquad \beta_5 = \frac{M_6}{M_1 M_4}\,, \qquad \beta_6 = \frac{M_4}{M_6}\,. \quad (4.4)$$

Nous avons aussi utilisé les projections verticales et horizontales basées sur les travaux de Ammar ([11]). Pour mesurer et comparer les projections entre une signature et un ensemble de références de signatures nous calculons deux paramètres *HF* et *VF* définis par les équations suivantes :

$$HF = \sum_{i=1}^{M} (HP_u(i) \times HRP(i))^{\frac{1}{2}},$$

$$VF = \sum_{j=1}^{N} (VP_u(j) \times VRP(j))^{\frac{1}{2}}, \qquad (4.5)$$

où :

- HP_u *et* VP_u sont les projections de la signature;
- M et N sont les dimensions de la signature;
- *HRP et VRP* sont les valeurs représentatives des projections des signatures de références.

Les valeurs HP_u , VP_u , *HRP et VRP* sont calculées pour une image donnée g(i,j) de dimension (N x M) à partir des équations suivantes :

$$HP_u(i) = \sum_{j=1}^{N} g(i,j), \qquad (1 \le i \le M) \qquad (4.6)$$

$$VP_u(j) = \sum_{i=1}^{M} g(i,j), \qquad (1 \le j \le N) \qquad (4.7)$$

$$HRP(i) = \frac{1}{m} \sum_{k=1}^{m} HP(i), \qquad (4.8)$$

$$VRP(j) = \frac{1}{m} \sum_{k=1}^{m} HP(j), \qquad (4.9)$$

où m est le nombre de signatures de références pour un signataire donné.

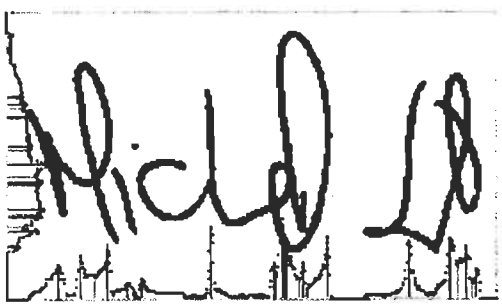

Figure 4.1 : Projections horizontales et verticales.

La hauteur et la largeur d'une signature sont aussi utilisées et leur calcul s'effectue à l'aide d'un algorithme rapide et simple de recherche des points noirs extrêmes de la signature et qui définissent les quatre bornes encadrant la signature. Cet algorithme est le même que celui utilisé au paragraphe 3.3.3. En effet, les dimensions de la signature sont égales aux dimensions de la fenêtre qui encadre la signature. La figure suivante montre le balayage d'une image pour le calcul de la largeur de la signature.

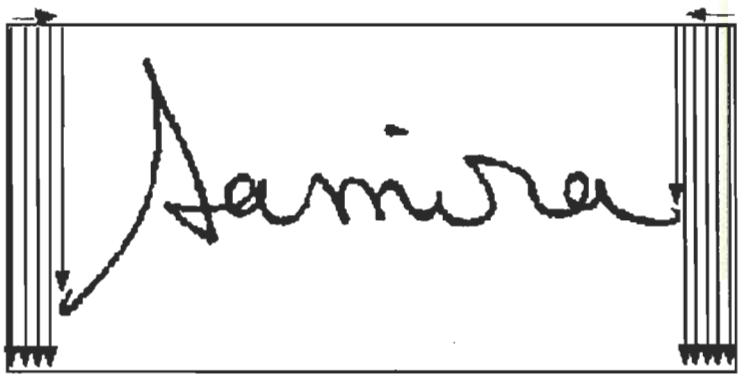

Figure 4.2 : Balayage de l'image.

Cependant, cet algorithme est peu robuste au bruit. En effet, la présence de pixels noirs parasites sur le fond fausse les calculs, car ces derniers seront considérés comme appartenant à la signature. Ce problème est partiellement résolu lors des étapes de pré-traitements ultérieures.

Dès que la hauteur et la largeur d'une signature sont calculées, le ratio=Hauteur/Largeur sera utilisé comme une mesure de la signature.

On appelle « partie d'une signature » tout sous-ensemble connexe de pixels noirs appartenant à la signature.

Après étude de quelques échantillons de signatures nous remarquons que les signatures de type américain se composent généralement de plusieurs parties correspondant, par exemple, au nom et au prénom du signataire avec un espace entre les deux. Par contre, une signature de type européen est généralement composée d'une seule partie connexe. Nous pouvons donc considérer le calcul du nombre de parties d'une signature comme une mesure discriminante des types de signatures.

Pour calculer le nombre de parties d'une signature S, nous procédons comme suit : nous calculons dans une première étape la projection horizontale de la signature à l'aide de l'équation (4.6) :

$$HP_S(i) = \sum_{j=1}^{N} g(i,j), \qquad (1 \le i \le M).$$

Nous parcourons ensuite $HP_S(i)$ pour i allant de 1 à M. Une signature est alors considérée comme contenant plusieurs parties si nous trouvons un intervalle d'entiers $[a,b] \subset [1,M]$ tel que :

$$\begin{cases} \forall i \in [a,b], \quad HP_S(i) = 0, \\ (b-a) \ge s, \end{cases}$$

où s est un seuil déterminé empiriquement en fonction de la largeur de la signature. Pour cette étude nous avons posé s égale à 5% de la largeur de l'image. La figure suivante montre un exemple de calcul de cette mesure.

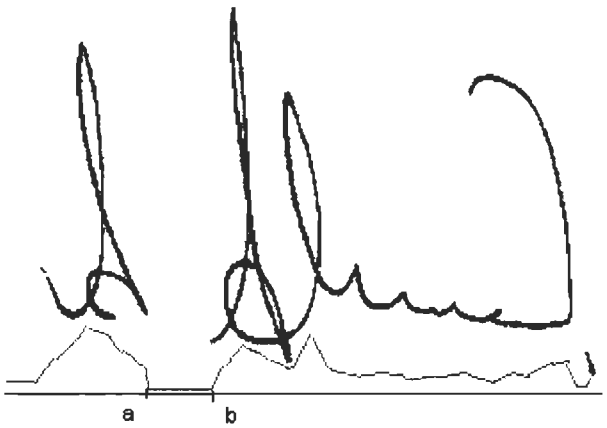

Figure 4.3 : Le nombre de parties de cette signature est 2.

D'autres mesures ont été testées, en l'occurrence l'enveloppe de la signature, la ligne de base et le squelette de la signature. Mais nous avons constaté que ces mesures ne fournissent aucune information sur le type de la signature, et conséquemment elles n'ont pas été retenues dans le système utilisé. Cependant nous les utiliserons dans le système de vérification des signatures.

Une fois les mesures citées ci-dessus calculées, la distance euclidienne suivante est utilisée pour discriminer le type de la signature,

$$distance = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{(F_i - \mu_i)_2}{\sigma_i^2} \right)^{1/2},$$

où F_i est la ième mesure de la signature, μ_i et σ_i sont la moyenne et la déviation de la i-ième mesure, n étant le nombre de signatures de référence.

Nous avons testé notre système sur un échantillon de 2225 signatures. Nous obtenons un taux d'erreur satisfaisant de 0.15% pour les signatures américaines et 3% pour les signatures européennes. Nous commenterons plus longuement ces résultats dans le paragraphe 5.1.

4.2 Système de vérification de la signature

Dans cette section nous présentons quelques définitions reliées au concept d'espace prétopologique qui serviront à mieux comprendre les mesures utilisées dans la mise en œuvre de la méthode de vérification des signatures. Nous décrirons aussi le module d'authentification qui utilise les réseaux de neurones. Ce système a été implémenté à l'aide du langage VC++6 sous Windows et il utilise le progilciel Matlab. Le module d'extraction des mesures est une application de type MDI (Multiple Document Interface) et dispose d'une interface ergonomique facile d'utilisation. Ce module produit un fichier texte qui sera utilisé par les modules d'apprentissage et de vérification.

4.2.1 Définition d'un espace Prétopologique

Soit E un ensemble non vide. On appelle fonction d'adhérence sur E toute application $a(.)$ *définie de* $\wp(E)$ *dans* $\wp(E)$ *vérifiant les deux propriétés suivantes :*

(i) $a(\phi)=\phi$,

(ii) $\forall A \in \wp(E)$, $A \subset a(A)$,

où $\wp(E)$ *désigne l'ensemble des parties de E.*

On dit alors que (E,a) est un espace prétopologique (M.Archoun [45]).

4.2.2 Extraction des caractéristiques

Le premier paramètre utilisé par notre système pour la vérification de signatures manuscrites est une (pseudo-distance) introduite par Elyassa et al ([21]). Cetts pseudo-distance est définie comme suit : soit (E,a) un espace pré-topologique, A et B deux sous-ensembles de E :

(i) $d(A,B) = \inf(k \mid B \subset a^k(A) \ \ et \ \ A \subset a^k(B))$, si $A \neq \phi$ et $B \neq \phi$,

(ii) $d(A,B) = +\infty$, si A= ϕ ou B= ϕ,

où $a^k(.)$ désigne l'adhérence d'ordre k .

Cet indice donne le nombre minimal de dilatations à appliquer aux ensembles A et B afin que le dilaté de A contienne B et réciproquement. Pour deux pixels, cette distance mesure la longueur du chemin le plus court les reliant.

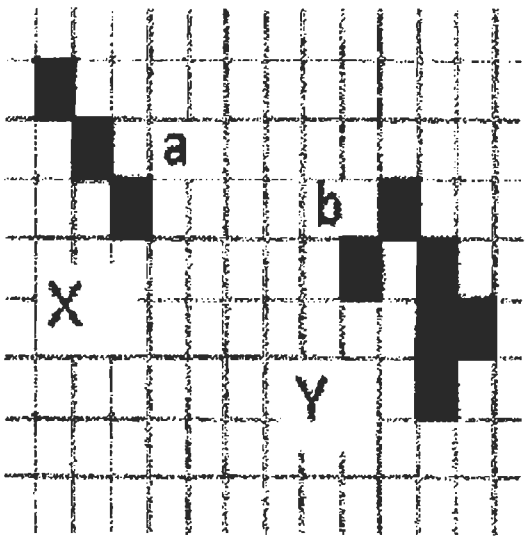

Figure 4.4 : d({a},{b})=6; d({a},X)=2;d({a},Y);d(X,Y)=11.

Elyassa et al ([21]) utilisent cette pseudo-distance pour mesurer la stabilité d'une classe de signatures. En effet, la distance entre deux signatures semblables est petite et s'annule si les deux signatures sont identiques. Ainsi, dans une même classe, l'écart des distances est réduit.

En pratique Elyassa et al ([21]) utilise l'inclusion partielle ($\varepsilon - inclusion$) pour le calcul des distances. Ceci permet de réduire le nombre de dilatations à appliquer étant donné que si ce nombre devient grand nous perdons toute information sur la forme de la signature. La valeur de ε est calculée de manière expérimentale (exemple : Inclusion à 90%; ε =0,10).

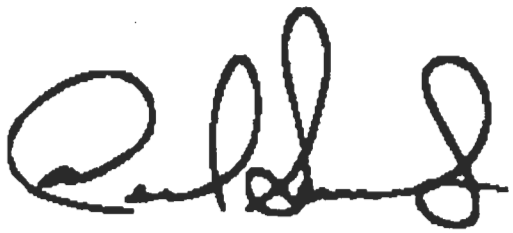

Figure 4.5: Signature S.

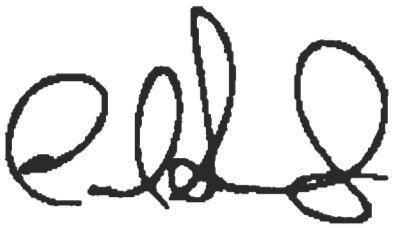

Figure 4.6: Signature S1
d(S1,S)=20.

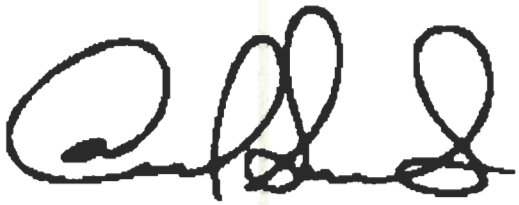

Figure 4.7: Signature S2
d(S2,S)=6, d(S1,S2)=22.

Les figures 4.5, 4.6, 4.7 montrent les distances obtenues entres trois signatures différentes. La signature S et S2 appartiennent au même signataire alors que la signature S1 est une imitation de la signature S. Nous remarquons que la distance entre S et S2 est petite comparativement aux distances entre S1 et S, et S1 et S2.

En plus des paramètres décrits au paragraphe 4.1 à l'étape d'identification du type de signature, nous utiliserons, les paramètres suivants :

- l'enveloppe de la signature,
- la ligne de base de la signature.

L'enveloppe supérieure (respectivement inférieure) de la signature est définie par l'ensemble des points supérieurs (respectivement inférieurs) du tracé binarisé de la signature.

L'algorithme d'extraction développé par Lee et al ([29]) consiste à balayer l'image colonne par colonne. Pour chaque colonne de l'image, nous gardons le premier point noir rencontré.

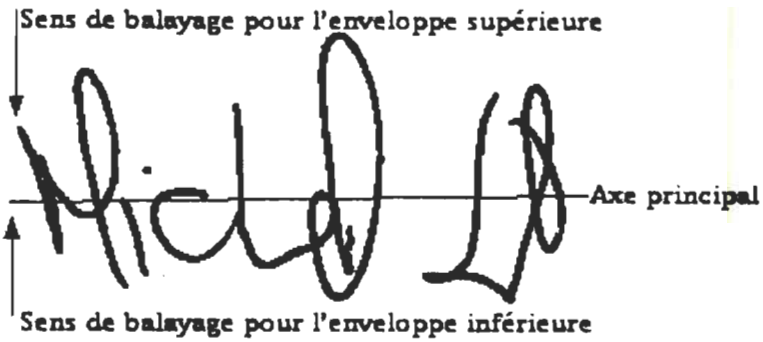

Figure 4.8 : Signature avec son axe principal.

Nous obtenons ainsi deux enveloppes, une supérieure et une inférieure comme le montre la figure ci-dessous.

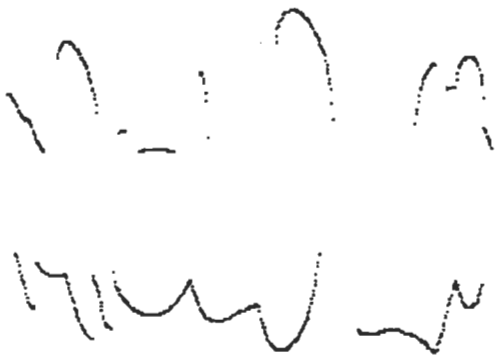

Figure 4.9 : Enveloppe supérieure et inférieure de la signature (4. 8).

La ligne de base d'une signature est une droite qui passe par le centre de la signature est qui suit son inclinaison. Nous avons utilisé la méthode introduite par Lee et al ([29]) qui permet de trouver l'inclinaison globale d'une signature à l'aide des vecteurs propres de la matrice de covariance des pixels noirs de la signature. Nous commençons par exprimer les coordonnées des pixels noirs de la signature dans un nouveau repère d'axes parallèles au repère de l'image. Ce nouveau repère est centré, au centre de gravité des pixels noirs de l'image que nous appellerons centroïde.

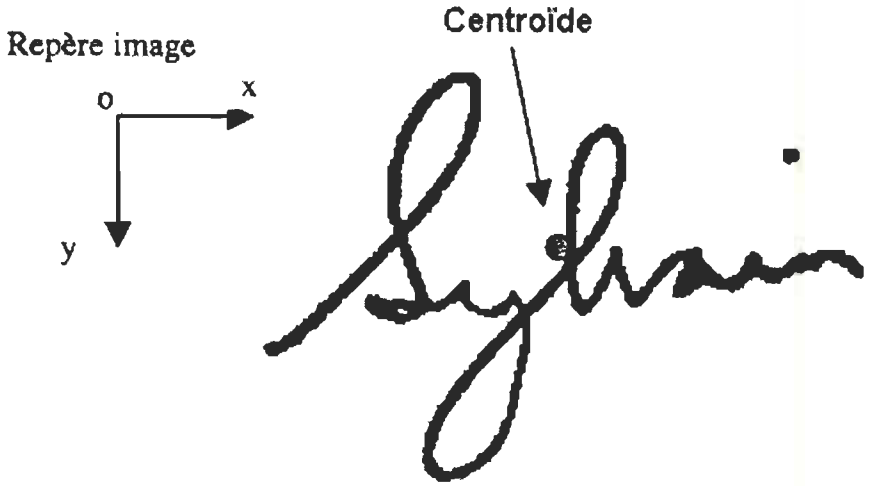

Figure 4.10 : Repère image et centroïde du motif.

Les coordonnées du centroïde sont données par les expressions :

$$\begin{cases} x\_\text{centroïd} = \sum_{i=1}^{n} \frac{x_i}{n}, \\ y\_\text{centroïd} = \sum_{i=1}^{n} \frac{y_i}{n}, \end{cases}$$

où n est le nombre de pixels noirs dans l'image et (x_i, y_i) sont les coordonnées des pixels par rapport au repère de l'image. Nous exprimons ensuite les coordonnées de tous les pixels dans le nouveau repère comme suit :

$$v_i = \begin{cases} x_{i\,nouveau} = x_{i\,ancien} - x\_\text{centroïd} \\ y_{i\,nouveau} = y_{i\,ancien} - y\_\text{centroïd} \end{cases}, i \in \{1,..,n\}.$$

En posant $V = (v_1, v_2, ..., v_n)$ la matrice de tous les pixels noirs de la signature, nous obtenons la matrice de covariance de nouveaux pixels d'après la formule suivante :

$$\sum = V'.V = \textit{matrice symétrique}\ (2 \times 2)$$

Pour cette matrice, nous obtenons deux valeurs propres et deux vecteurs propres. Nous retiendrons le vecteur ayant la valeur propre maximale. L'orientation de ce vecteur correspond à l'inclinaison globale de la signature.

Pour trouver la ligne de base nous déplaçons la ligne colinéaire au vecteur propre sur la signature jusqu'à la position où la ligne rencontre un maximum de pixels noirs.

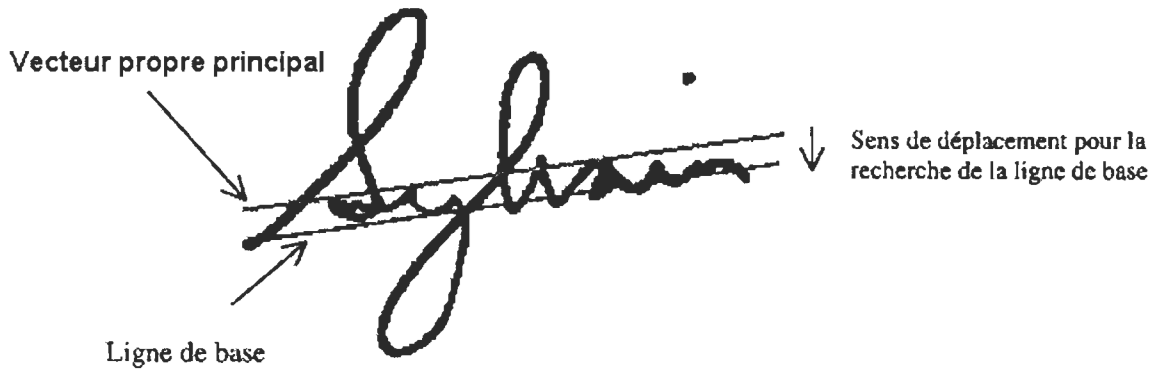

Figure 4.11 : Recherche de la ligne de base.

La figure 4.12 montre des exemples de signatures et de leur ligne de base calculée par cette méthode. Cette mesure donne des informations pertinentes sur l'habitude du scripteur mais ne représente aucun intérêt si le signataire change souvent l'inclinaison de sa signature.

a.

b.

c.

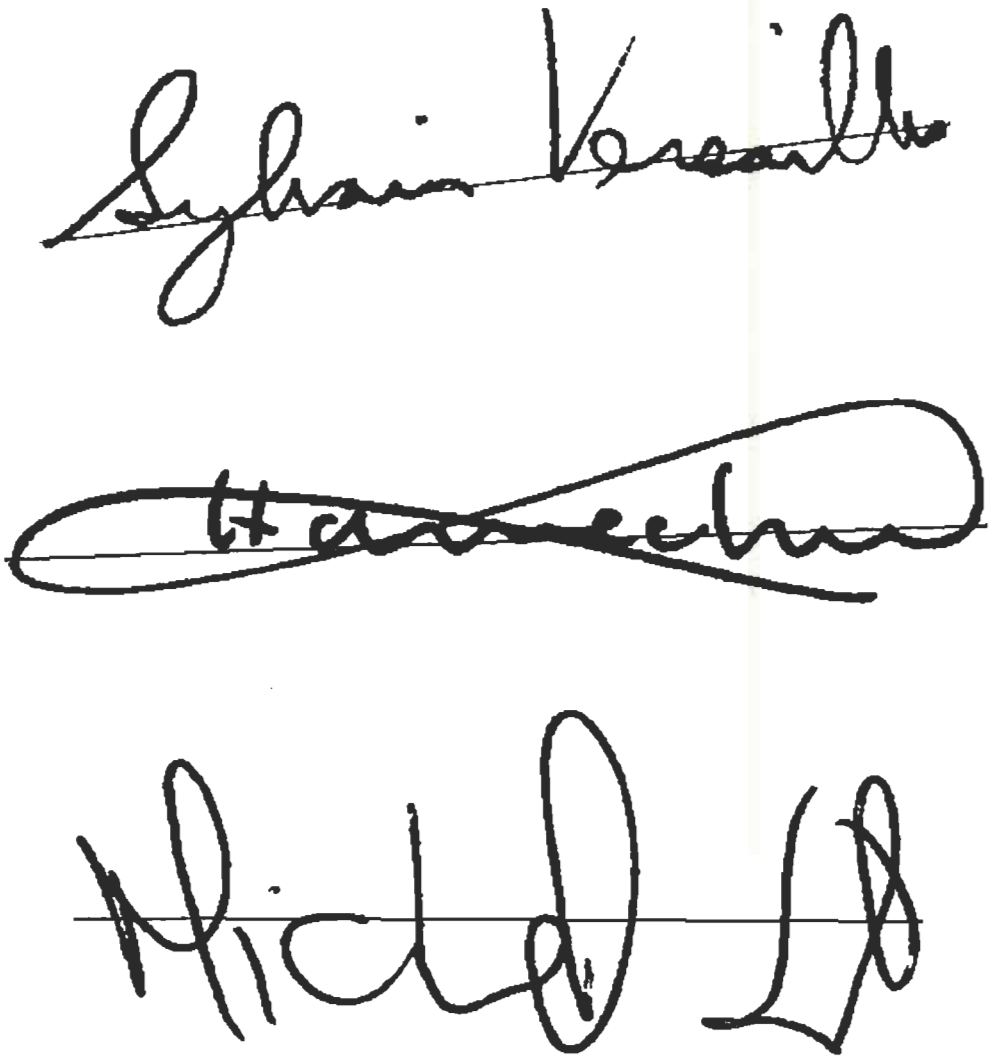

Figure 4.12 : Exemples de signatures et de leur ligne de base.

4.2.2 Réseaux de neurones

Dans cette section nous décrivons le système d'authentification des signatures manuscrites faisant l'objet de ce mémoire. Ce système est basé sur l'architecture des réseaux de neurones et utilise les caractéristiques et mesures décrites précédemment dans le but d'authentifier une signature donnée.

Nous avons utilisé un réseau de type "aller-retour" (back propagation) où les neurones collaborent dans le sens "aller" en réalisant un traitement de plus en plus élaboré de l'information et dans le sens "retour", en permettant la correction des traitements antérieurs.

Ces réseaux présentent l'avantage de corriger les poids des nœuds du réseau en continue afin d'améliorer les résultats obtenus. Cependant, ces corrections risquent de les rendre inutilisables s'ils ne sont pas bien entraînés. En effet, une erreur dans la reconnaissance d'une signature risque de fausser la correction des poids du réseau. Cette erreur aura donc, des répercussions sur tout les calculs subséquents.

L'algorithme d'authentification utilise deux réseaux distincts en aval. Un premier réseau utilise l'information fournie par la distance prétopologique pour décider de l'authenticité d'une signature. Ce réseau fournit comme résultat une valeur comprise entre 0 et 1. Une signature est considérée authentique si la valeur retournée par le réseau est supérieure à s_1=0.8 et est considérée comme fausse si la valeur retournée est inférieure à s_2=0.4. En cas d'indécision (la valeur retournée par le réseau est entre s_1 et s_2), un deuxième réseau utilise les mesures géométriques, énumérées à la section précédente, pour statuer sur l'authenticité de la signature. Nous avons choisi les valeurs de s_1 et de s_2 par expérience (essai-erreur) de façon à maximiser les performances du système (TVR et TFA faibles).

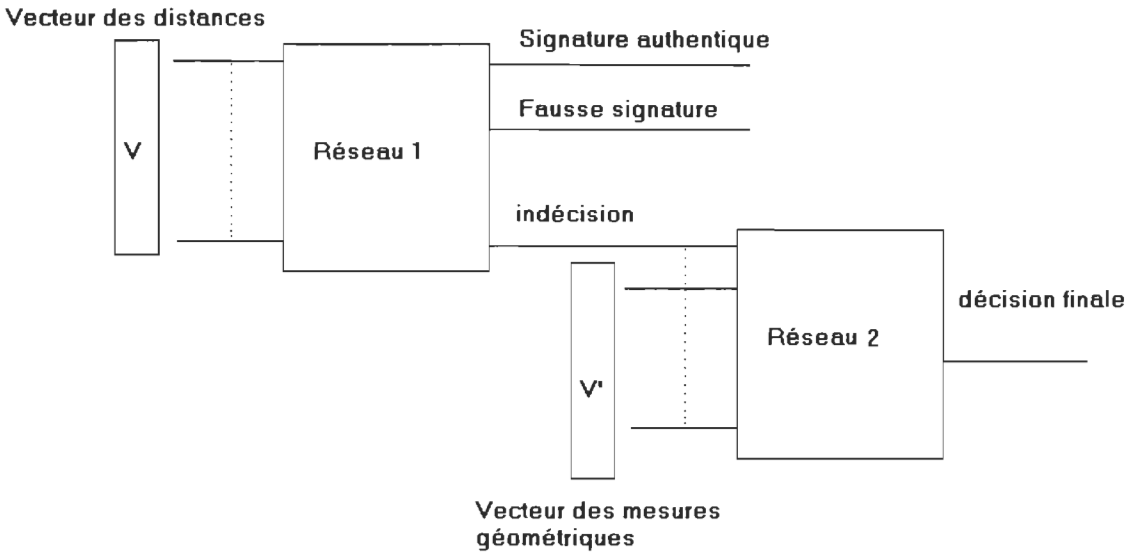

Figure 4.13 : Architecture des réseaux.

Tel que mentionné à la section 3.1 nous partageons la base de données en deux sous-ensembles P_1 *et* P_2 ; la partie P_1 étant utilisée pour l'apprentissage du système alors que P_2 est utilisée lors de la phase expérimentale.

Phase d'entraînement

Lors de cette étape, les signatures de P_1 sont utilisées dans la construction des vecteurs V et V' (figure 4.13). Le vecteur V contient les distances prétopologiques croisées intra-classes entre les signatures et le vecteur V' contient les mesures géométriques des signatures de P_1.

Pour chaque signataire s, m signatures authentiques et n fausses signatures sont sélectionnées. Pour chaque signature i de cette sélection nous calculons les distances d_{ij},

$$d_{ij} = d(S_i, S_j) \quad 1 \le j \le m+n ;$$

et on note V^i le vecteur associé à la signature i, pour $1 \le i \le m+n$;

$$V^i = \begin{pmatrix} d_{i1} \\ d_{i2} \\ . \\ d_{ij} \\ . \\ d_{i(n+m)} \end{pmatrix}, \quad 1 \le j \le n+m.$$

Afin d'entraîner le réseau 1 (figure 4.13) nous lui soumettons l'ensemble des vecteurs V^i, i=1,...,m+n .

Pour chaque signataire s, l'ensemble des (m+n) signatures sélectionnées est utilisé pour calculer les vecteurs V^{ii} $(1 \leq i \leq m+n)$ des mesures géométriques. Pour chaque signature i le vecteur V^{ii} s'écrit:

$$V^{ii} = \begin{pmatrix} m_1^i \\ . \\ . \\ . \\ m_p^i \end{pmatrix},$$

où $m_k^i, (1 \leq k \leq p)$ est la k-ième mesure géométriques (voir chapitre 3 et 4) et p est le nombre total de ces mesures, soit p=7 (mesure des moments, projection horizontale, projection verticale, ratio hauteur/largeur, nombre de parties de la signature, enveloppe, ligne de base).

Pour entraîner le réseau 2 (figure 4.13) nous lui soumettons donc les vecteurs V^{ii} avec $1 \leq i \leq m+n$.

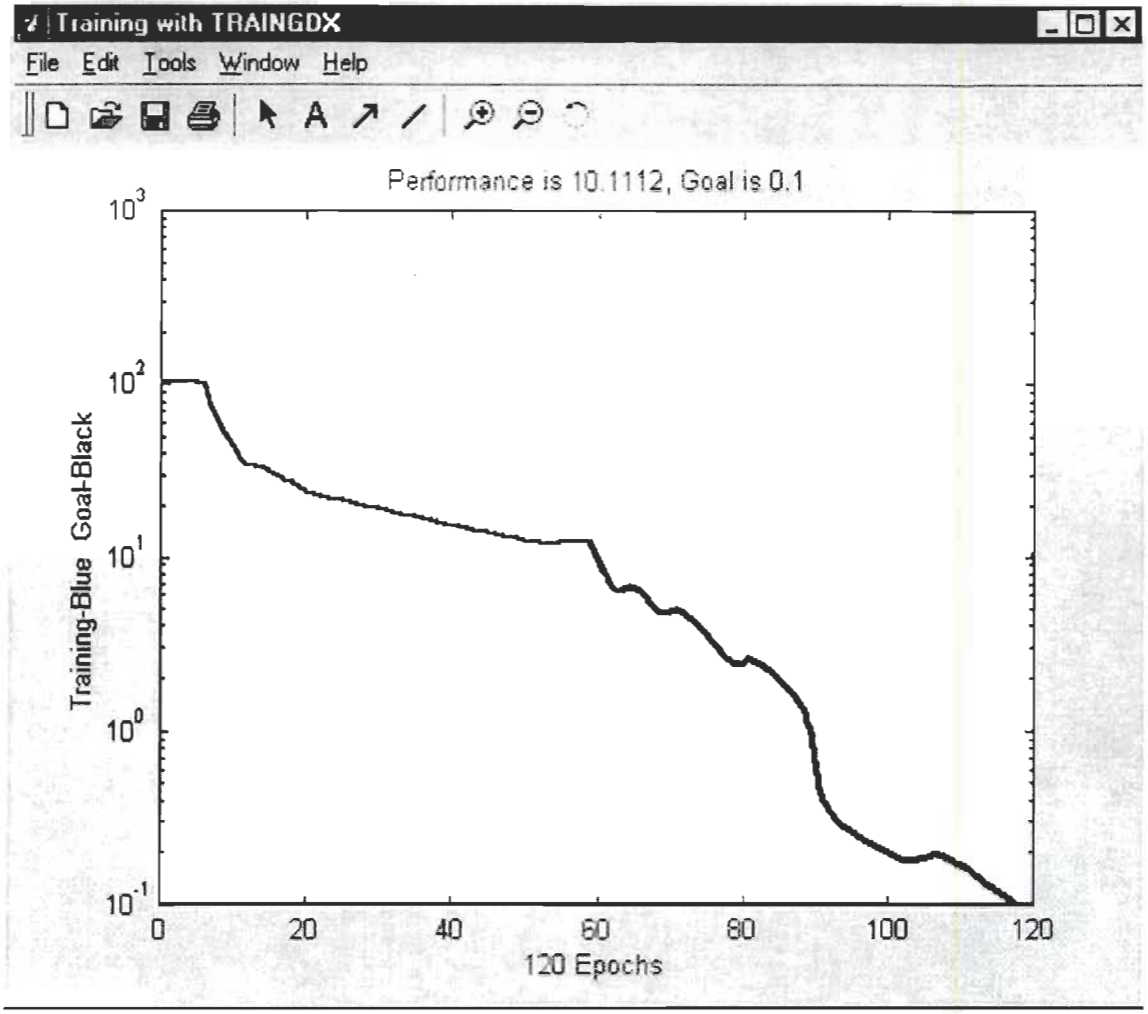

Figure 4.14 : Phase d'entraînement des réseaux de neurones par Matlab.

La figure 4.14 illustre la phase d'entraînement du système sous Matlab. L'axe des X représente le nombre d'itérations (Epochs) que le réseau exécute pour corriger ses poids. L'axe Y représente les performances du système en terme d'erreur. Dans cet exemple l'erreur désirée a été fixée à 0.1 et le système a effectué 120 itérations pour atteindre ce seuil. Idéalement, un réseau performant devrait avoir une erreur égale à 0, mais en pratique une erreur de 0.01 est considérée comme étant satisfaisante.

Chapitre 5

Simulations et résultats

Dans ce chapitre, nous décrivons les résultats obtenus lors de la phase des tests des systèmes d'authentification et de la reconnaissance du type de la signature. De plus nous discuterons des améliorations que nous pouvons apporter au système développé dans le cadre de cette recherche.

5.1 Résultats : système de reconnaissance du type

Nous avons testé le système proposé sur un échantillon de 2225 signatures réparties en deux classes : signatures européennes et signatures américaines. Nous présentons au système une seule signature à la fois sans condition préalable sur son type et nous calculons le nombre d'échecs. Le taux d'erreur est obtenu en divisant le nombre d'échecs par le nombre total des signatures. Le tableau suivant montre les résultats obtenus pour chacune des classes .

Type de signature	Taux d'erreur
Signatures américaines	0.15%
Signatures européennes	3%

Tableau 5.1 : Taux d'erreur pour les signatures de type américain et européen.

Les figures (5.1, 5.2 et 5.3) montrent trois exemples de signatures de différents types et les résultats obtenus.

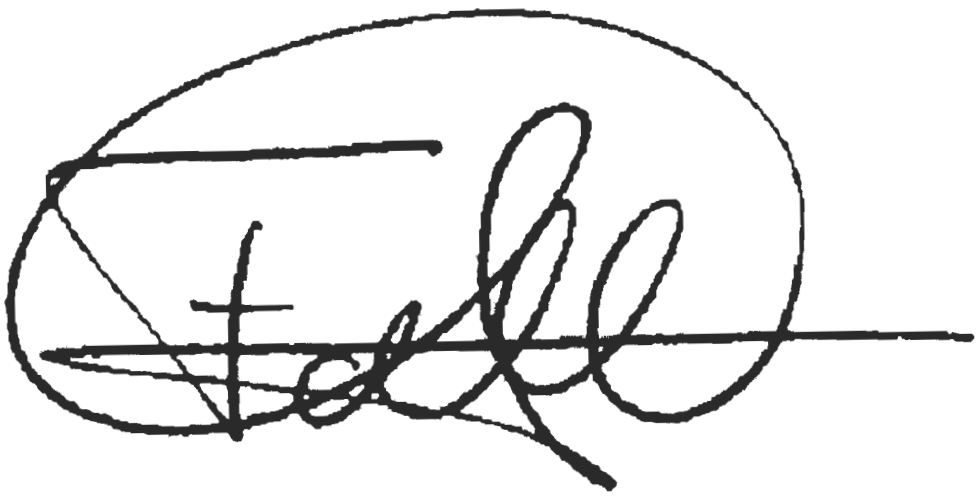

Figure 5.1 : Signature européenne reconnue avec succès.

Figure 5.2 : Signature américaine reconnue avec succès.

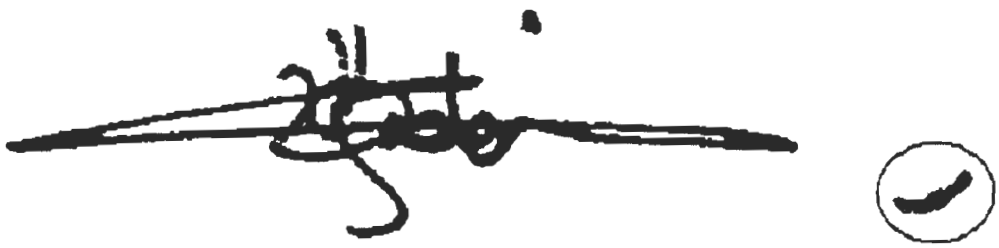

Figure 5.3 : Signature européenne reconnue comme signature américaine.

Malgré les résultats obtenus, l'algorithme de reconnaissance reste encore très sensible au bruit et présente quelques lacunes auxquelles il faut remédier. C'est le cas, par exemple, de la signature à la figure 5.3 où la présence de bruit dans l'image fausse le résultats. Une façon simple de résoudre ce problème serait d'ajouter des filtres de prétraitements pour éliminer le bruit, mais ceci risque de détériorer la qualité de l'image.

Une autre méthode, consisterait à ajouter une condition pour ne tenir compte que des parties de la signature qui ont une largeur importante relativement à la largeur de la signature entière. Ceci permettrait à l'algorithme d'ignorer les parties les plus petites de la signature qu'il associerait alors à du bruit.

5.2 Résultat du système de vérification

Pendant la phase d'apprentissage nous définissons pour chaque signataire j une classe C_j à partir de m signatures authentiques et n fausses signatures. Pour tester les réseaux de neurones nous utilisons l'ensemble P_2 définie dans 4.2.2. Pour chaque signature de cet ensemble nous calculons sa distance prétopologique avec les (m+n) signatures pour construire le vecteur V. Nous présentons ensuite ce vecteur au réseau 1. En cas d'indécision, nous calculons les mesures géométriques pour construire les vecteur V' que nous présentons au réseau 2.

Cette méthode donne de meilleurs résultats que ceux présentés par Elyassa *et al* ([21]) qui construit les classes représentatives de chaque signataire à partir du calcul de la moyenne et de l'écart type. Cependant, les performances du système dépendent grandement du choix de m et n pour constituer la classe représentative du signataire. La figure suivante montre les variations des NFA[8] et NVR[9] dépendamment des variations de m et n pour un signataire fixé.

Valeur de m	valeur de n	n=2	n=6	n=10	n=14	n=18	n=20
m=2	NFA	30	25	20	18	25	13
	NVR	25	20	23	15	17	14
m=6	NFA	26	29	23	25	20	15
	NVR	22	18	20	16	13	10
m=10	NFA	20	16	22	13	10	6
	NVR	19	15	13	8	12	11
m=14	NFA	28	20	19	17	10	7
	NVR	14	12	10	13	9	9
m=18	NFA	32	17	13	14	10	8
	NVR	8	5	7	4	2	5
m=20	NFA	24	15	10	7	4	1
	NVR	6	7	2	5	4	3

Tableau 5.2 : Variations des NFA et NVR dépendamment des variations de m et n pour un signataire donné.

[8] Nombre des faux acceptées
[9] Nombre des vrais rejetées

Pour un signataire donné nous avons donc testé notre méthode de vérification plusieurs fois, en changeant à chaque test le nombre des signatures présentées au réseau pendant l'étape d'apprentissage. Nous remarquons alors qu'en augmentant les nombres m et n simultanément le système devient de plus en plus précis. Ceci s'explique par le fait que le système dispose de plus d'informations lors de la phase d'entraînement et réussi à mieux différencier les signatures authentiques des fausses.

Malheureusement, le nombre réduit de signatures disponibles pour chaque signataire ne nous permet pas de présenter au réseau plus que 40 signatures pendant l'étape d'apprentissage. Il serait intéressant de comparer ces résultats avec les performances du système avec un nombre plus élevé de signatures.

En regard des résultats de ces tests, il a été décidé de fixer pour tout les signataires m=20 et n=20 et de tester le système pour comparer nos résultats avec ceux obtenus par Elyassa et al ([21]).

A cet effet , nous avons divisé le sous ensemble P_2 est divisé en 20 classes distinctes, chacune représentant un signataire différent que nous notons : classe a, classe b,..., classe t.

Pour chaque classe nous comparons le nombre des vrais rejetés (NVR) et le nombre des faux acceptés (NFA) obtenus par les deux méthodes. Nous obtenons les résultats présentés au tableau 5.3:

Classe	Calcul de moyenne		Réseaux de neurones	
	NVR	NFA	NVR	NFA
a	0	0	0	0
b	1	0	0	0
c	3	0	2	0
d	0	0	0	0
e	1	0	1	0
f	0	0	0	1
g	0	0	0	0
h	3	0	1	0
i	0	0	0	0
j	0	0	0	0
k	0	0	0	0

Tableau 5.3 : Tableau comparatif des résultats obtenus pour chacune des deux méthodes pour les classes a..k.

Classe	Calcul de moyenne		Réseaux de neurones	
	NVR	NFA	NVR	NFA
l	0	0	0	0
m	1	0	0	0
n	2	0	1	1
o	1	0	0	0
p	2	6	1	3
q	2	0	2	0
r	0	0	0	0
s	0	1	0	0
t	1	0	0	0

Tableau 5.4 : Tableau comparatif des résultats obtenus pour chacune des deux méthodes pour les classes l..t.

Les taux TFA et TVR ont été calculées pour chaque méthode.

	TVR	TFA
Calcul de moyenne	4%	1,75%
réseaux de neurones	2%	1,25%

Tableau 5.5 : Les taux d'erreurs .

Le système développé par Elayssa et al ([21]) utilise comme critère d'agrégation une méthode statistique simple basée sur le calcul des moyennes. Ce système présente de bon résultats s'il dispose d'un très grand nombre de signatures de référence, ce qui est malheureusement impossible dans le cas d'une utilisation réelle de ce système. Les réseaux de neurones permettent de remédier à ce problème, puisque le système décrit dans ce mémoire donne de meilleurs résultats avec un nombre restreint de signatures de référence.

L'utilisation de la distance prétopologique comme caractéristique pour les signatures permet de décrire les habitudes de chaque signataire. Elle représente une mesure locale permettant de distinguer une signature véritable d'une contrefaçon habile semblable au niveau de la forme à l'original. Les caractéristiques géométriques fournissent quand à elles, une mesure globale pour différencier les faux grossiers[10] des signatures authentiques.

[10] Les faux grossiers sont des faux qui ne ressemble pas au niveau de la forme à un original.

Chapitre 6

Conclusion

Dans ce mémoire, nous avons développé deux méthodologies complémentaires. Une première permet de créer un système automatique de reconnaissance du type des signatures manuscrites, alors que la deuxième est une approche d'authentification des signatures manuscrites se basant sur les réseaux de neurones et utilisant des mesures géométriques et une distance prétopologique.

Le premier système de reconnaissance utilise des mesures géométriques pour caractériser les signatures. L' «ensemble des signatures manuscrites » est partagé en deux classes : *signature de type américain* et *signature de type européen.* Ce système est constitué de trois étapes. La première étape est basée sur le prétraitement des images et a pour but de réduire le bruit qui s'introduit lors de la saisie et de la numérisation des images. Les images traitées ainsi sont ensuite transmises à l'étape d'extraction des caractéristiques. Dans cette étape, l'algorithme extrait les mesures géométriques de l'image de la signature (projections, moments, parties d'une signature...). La troisième étape correspond au calcul de la distance entre le vecteur des caractéristiques de la signature étudiée et les vecteurs moyens de chaque classe et fournit une décision sur le type de la signature.

Ce système à été validé sur plus de 2000 signatures et les résultats obtenus sont satisfaisants. En effet, le système tel que proposé, présente de bonnes performances tant au niveau de la vitesse d'exécution qu'au niveau de la précision des résultats.

Le deuxième système développé est un système de vérification de la signature. Il permet, grâce à une mesure de distance prétopologique développée par Elyassa et al ([21]), de vérifier si la signature traitée est une signature authentique ou non. La contribution de notre approche est surtout au niveau de l'authentification où notre système utilise un réseau de neurones 'aller-retour' pour la vérification des signatures.

Nous avons montré qu'en utilisant des mesures géométriques globales et une distance prétopologique, la fiabilité du système en regard du traitement des faux grossiers a augmenté. Ceci nous permet d'avoir de meilleurs résultats comparativement aux autres méthodes utilisant la même distance prétopologique. Nous avons testé notre système sur 400 signatures réparties en 20 classes et obtenu des taux TFA=2% et TVR=1.25% au lieu de TFA=1,75% et TVR=4,00% qui sont obtenus par Elyassa et al ([21]).

Ces résultats sont meilleurs que ceux obtenus par la plupart des autres méthodes utilisant des réseaux de neurones citées dans le chapitre 2. En effet, le système proposé par Huang et al. [13] donne des taux TFA=11.8% et TVR=11.1% alors que celui de Fadhel et al ([18]) produit des taux TFA=6.2% et TVR=5.5%. Cependant, la comparaison entres toutes ces méthodes est difficile étant donné la différence des bases de données et des types de faux utilisés. Ainsi, Huang et al. [13] utilisent une base de 3528 signatures dont 144 sont des faux par imitation, Fadhel et al ([18]) utilisent une base de 300 signatures produites par 30 individus et des faux grossiers pour valider leur système alors que notre base de donnée est constituée de 800 signatures et tous les faux sont de type «grossier».

Bibliographie

[1] J.Gayet, "Manuel de Police Scientifique", Payot, Paris, 1961.

[2] J.Mathyer, "The Expert Examination of Signature", Journal of criminal law, Criminology and police science, Vol.5, N° 3, Mai-Juin 1961.

[3] W.R.Harrison, "suspect Documents, Their Scientific Examination", Chicago, Nelson-Hall publisher, 1981.

[4] R"Saferstein, "Forensic Science Handbook", Prentice-Hall, 1982.

[5] I.W.Evett et R.N.Totty, "A Study of Variation in the Dimensions of Genuine Signatures", Journal of the forensic Science Society, Vol.25, pp 205-215, 1985.

[6] E.Locard, "Les Faux en Écriture et leur Expertise", Payot, Paris 1959.

[7] S.A.Slyter, "Forensic Signature Examination", Charles C Thomas publisher,1995.

[8] D.j.Burr, "Experiments on Neural Net Recognition of Spoken and Written Text." IEEE Transacion on ASSP, Vol 36 N° 71988.

[9] R.Sabourin, M.Cheriet, G.Genest, "An extented-shadow-code based approach for off-line signature verification. Proceedings of the second International Conference on Document Analysis and Recognition", Vol 29 N°3, 1996.

[10] M.Ammar, Y.Yoshiba, T.Fukurama, "Structural description and classification of signature images". Pattern Recognition, Vol 23,N°7, 1990.

[11] M.Ammar, "Performance of parametric and reference pattern based features in static signature verification:a comparative study", 10th International Conference on Pattern Recognition, Vol 1, 1990.

[12] E.R.Brocklehurst, "Computer methods of signature verification", National Phys. Lab, Report N° NPL-DITC-41/84,1984.

[13] K.Huang, H.Yan,"Off-Line Signature Verification based on geometric feature extraction and neural network classification", Pattern Recognition, Vol 30 ,N°1, 1997.

[14] R.Sabourin, G.Genest,F.Prêteux,"Off-Line Signature Verification by Local Granulometric Size Distributions",IEEE Transactions On Pattern Analysis and Machine Intelligence, VOL 19,N°9,1997.

[15] Y.QI, R.Hunt, "Signature Verification Using Global and Grid Features", Pattern Recognition, VOL 27, N°12, 1994.

[16] Y.QI, R.Hunt, "A Multiresolution approach to computer verification of handwritten signatures", IEEE Transactions on Image Processing Vol 4, N°6, 1995.

[17] C.Sansone, M.Vento, "Signature Verification: Increasing performance by using a Multi-Stage System", Pattern Analysis & Application Vol3, 2000.

[18] E.A. Fadhel at P.Bhataacharyya, "Application of a steerable wavelet transform using neural network for signature verification", Pattern Analysis & Applications, N° 2, 1999

[19] H.Baltzakis,N.Papamarkos, "A new signature Verification technique based on a two-stage neural network classifier", Engineering Applications of Artificial Intelligence,n°14, 2001.pp. 95-103

[20] Y.Mizukami, H.Miike, M.Yoshimura, I.Yoshimura, "An off-line signature verification system Using an extracted displacement function", IEEE Transactions On Pattern Analysis and Machine Intelligence, 1999.

[21] M. Elyassa , D. Mammass and F. Nouboud, "Signature Verification Based on a Pretopological Approach", International Conference on Image and Signal Processing, Agadir, Morocco, 2001.

[22] M.Lamure, "Espace abstraits et reconnaissance de formes: Application aux traitements d'image digitale", Thèse de doctorat d'etat- UCB. Lyon I (France), 1987

[23] B.Fang, Y.Y.Wang, C.H.Leung, Y.Y.Tang, P.C.K.kwok, K.W.Tse, Y.K.Wong, "A smoothness index based approach for off-line signature verification", In Proceeding of the 5th International Conference On Document Analysis and Recognition, 1999.

[24] V.E Ramesh,M.Narasimha Murty, "Off-line signature verification using genetically optimized weighted features", Pattern Recognition, VOL 32, 1999,pp217-233.

[25] R.Bajaj,S.Chaudhury, "Signature Verification using multiple Neural Classifiers", Pattern Recognition,Vol.30,n°1, 1997.

[26] D.K.R McCormack. J.F. Pedersen, " Neural network signature verification using Haar wavelet and Fourier transforms",SPIE,Vol.2064,1993,pp 14-25.

[27] R.Plamondon,G.Lorette,"Automatic signature verification and writer identification-the state of the art", Pattern Recognition, VOL 22, N°2, 1989.

[28] S.Wilkinson, W.Goodman, "Slope Histogram Detection of forged handwritten signatures","High-Speed Inspection Architectures,Barcoding, and Character Recognition", 1990.

[29] S.Lee,C.Pan,"Off-line tracing and representation of signature", IEEE Transactions On Systems,Man, and Cybernetics, VOL 22, N°4, 1992.

[30] F. Nouboud and R. Plamondon, "Global Parameters and Curves for Off-Line Signature Verification", Int. Workshop on Frontiers in Handwriting Recognition, 1994, pp. 145-154.

[31] R.Sabourin,J-P.Drouhard,E.Sum-Wah,"Shape Matrices as a Mixed Factor for Off-Line Signature Verification", In Proceeding of the 4th International Conference On Document Analysis and Recognition, 1997.

[32] S.Djeziri, F.Nouboud, R.Plamondon,"Extraction of signature from check background based on a filiformity criterion", IEEE Transactions On Image Processing, VOL 7, N°10, 1998.

[33] H.Cardot,"Étude et spécification d'une architecture de réseaux neuronaux pour l'authentification de signatures manuscrites statiques" , Université de Caen, 1993.

[34] R.Sabourin, "Une approche de type compréhension de scène appliquée au problème de la vérification automatique de l'identité par l'image de la signature manuscrite",Université de Montréal,1990.

[35] J.P.Crettez, "Premier degré de caractérisation des écritures manuscrites:Essai de regroupement des écritures en familles", 1994.

[36] IW.Evett, RN.Totty, "A study of the variation in the dimension of genuine signature" , Forensic Science Society, 1985.

[37] J.P.Drouhard, R.Sabourin, M.Godbout , "Evaluation of training methode and of various rejection criteria for a neural network classifier used for off-line signature verification" , IEEE Int'l Conf. Neural Networks, Orlando,Fla June 26-July2 , 1994.

[38] J.P.Drouhard, R.Sabourin, M.Godbout, "A neural approach to off-line signature verification using directional PDF" , Pattern Recognition , VOL 29, N°3, 1996.

[39] M.Ammar and Y.Yoshida and T.Fukumara, "Off-line preprocessing and verification of signature" , Pattern Recognition and Artificial Intelligence, VOL 2, N°4, 1988.

[40] F.Nouboud, "Contribution à l'étude et à la mise au point d'un système d'authentification de signatures manuscrites", Université de Caen, 1988.

[41] F.Nouboud, A. Chalifour, " Definition of a Writing Pseudo-Order for Off-Line Handwritten Signature Verification ", International Conference on Image and Signal Processing, Agadir, Morocco, 2001.

[42] M.Hu, "Visual Pattern recognition by moment invariants", IEEE Transactions Inform, Theory, IT8:179-189,1962.

[43] R.C.Doria, E.C.B.C Fitho, P.J.L.Adeodato, "How Distorsions in Different Size Signatures Influence Feature Extraction Techniques", International Conference on Image and Signal Processing, Agadir, Morocco, 2001.

[44] S.Maitra, "Moment Invariant", Em IEEE 67, pp 697-699, 1979

[45] M.Archoun, " Modélisation prétopologique de la segmentation par croissance de régions des images à niveaux de gris", Université Claude Bernard Lyon I, 1993.